

# UNIVERZITA KARLOVA PŘÍRODOVĚDECKÁ FAKULTA

Studijní program: Molekulární a buněčná biologie, genetika a virologie

Studijní obor: Molekulární a buněčná biologie, genetika a virologie



**Mgr. Michal Sýkora**

Charakterizace nekanonické RNA polymerázy kódované lineárními  
plasmidy kvasinek

Characterization of non-canonical RNA polymerase encoded by the yeast linear  
plasmids

Disertační práce

Školitel: Mgr. Václav Vopálenský, Ph.D.

Praha, 2019

**Prohlášení:**

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 21.8. 2019

.....  
Michal Sýkora



## Poděkování

Na tomto místě bych rád poděkoval všem, bez kterých by tato práce nemohla vzniknout. Děkuji především mému školiteli Václavovi Vopálenskému za vše, co mě naučil, za pravidelnou diskuzi všech možných a nemožných hypotéz, a za všechny experimenty, které mi povolil uskutečnit, byť se v danou chvíli zdály sebedivočejší. Další dík patří ostatním současným a bývalým členům Laboratoře biochemie RNA, jmenovitě Josefu Novákovi, Silvii Mrvové, Anasovi Khawaja, Miroslavu Příbylovi a Martinu Cencialovi, díky kterým jsem se do laborky každý den těšil. Další poděkování patří kolegům z jiných laboratoří, kteří se příležitostně nechali zlákat mými nápady a investovali do nich trochu svého času a prostředků, jmenovitě Liborovi Krásnému, Petře Sudzinové, Haně Španielové, Janě Váňové a Zoře Mělkové. Poděkování patří i Natálii Suchánkové za vřelou a příkladnou laborantskou práci. Special thanks goes to prof. Finn Werner and all members of the RNAP lab at University College London (especially to Gwenny Cackett, Thomas Fouqueau, and Declan Barker) for making my Erasmus+ stay such a wonderful experience. Největší dík však patří mojí rodině a nejbližším za vytvoření příjemného zázemí pro bádání a za trpělivost a podporu během celého nekrátkého studia.

Tato práce byla vypracována v letech 2013–2019 za podpory grantového projektu P305/12/G034 Grantové agentury České republiky.

# ABSTRAKT

Transkripce je hlavním kontrolním bodem genové exprese. Tento proces závisí na proteinovém komplexu vícepodjednotkových RNA polymeráz, které jsou mimořádně konzervované mezi všemi buněčnými organismy. U mimochromozomálních dědičných elementů, jako jsou organely, víry a plasmidy, je transkripce závislá na RNA polymerázách buněk hostitelských organismů, nebo tyto elementy kódují RNA polymerázu vlastní. Předpokládaná nekanonická RNA polymeráza, skládající se ze dvou podjednotek, je kódována také lineárními cytoplasmatickými plasmidy kvasinky *Kluyveromyces lactis* a velmi pravděpodobně přepisuje pouze geny těchto plasmidů. Kromě dvou podjednotek RNA polymerázy kódují lineární plasmidy *Kluyveromyces lactis* ještě další dvě předpokládané komponenty transkripčního aparátu, a to *capping* enzym přidávající čepičku na 5' konce mRNA, a předpokládanou DExD/H box helikázu.

Charakterizace unikátního a nepříliš probádaného transkripčního aparátu plasmidů *Kluyveromyces lactis* byla hlavní náplní této práce. Cílem bylo především: 1) objasnit evoluční původ transkripčního aparátu lineárních plasmidů; 2) popsat architekturu transkripčního komplexu lineárních plasmidů *in vivo* se zaměřením na předpokládané vazebné partnery RNA polymerázy; 3) odhalit mechanismy iniciace a terminace transkripce kvasinkových lineárních plasmidů.

Hlavními výsledky této práce jsou zjištění, že: 1) nekanonická RNA polymeráza a promotory genů kvasinkových lineárních plasmidů mají původ blízký poxvirům; 2) podjednotky RNA polymerázy, mRNA *capping* enzym a předpokládaná helikáza tvoří komplex *in vivo*, přičemž předpokládaná helikáza specificky interaguje s plasmidovou DNA *in vivo*; 3) během iniciace transkripce dochází k 5' polyadenylaci transkriptů klouzáním RNA polymerázy na sekvenčním elementu promotoru a na tvorbě 3' konců transkriptů se podílí RNA vlásenky *in vivo*, pravděpodobně mechanismem vnitřní terminace transkripce.

**Klíčová slova:** *Kluyveromyces lactis*, lineární plasmidy, nekanonická RNA polymeráza, transkripce

# ABSTRACT

Transcription is the control point of gene expression. This process relies on protein complex of multisubunit RNA polymerases, which are extremely conserved among all cellular organisms. Transcription of extrachromosomal hereditary elements such as organelles, viruses and plasmids is dependent on host cellular RNA polymerases or intrinsic RNA polymerase is contained within these elements. Putative non-canonical two-subunit RNA polymerase is also encoded by linear cytoplasmic plasmids of the yeast *Kluyveromyces lactis* and most likely transcribes genes of these plasmids. Besides the two subunits of RNA polymerase encoded by linear plasmids of *Kluyveromyces lactis* there are another two estimated components of the transcription apparatus, namely capping enzyme that adds the cap to 5' mRNA ends and putative DExD/H box helicase.

Characterization of the unique and underexplored transcription machinery of *Kluyveromyces lactis* plasmids was the principal objective of this work. The main goal was to: 1) clarify evolutionary origin of the linear plasmid transcription apparatus; 2) describe architecture of the linear plasmid transcription complex *in vivo* focused on putative RNA polymerase binding partners; 3) reveal mechanisms of transcription initiation and termination of the yeast linear plasmids.

The main results of this work are the findings that: 1) non-canonical RNA polymerase and promoters of the yeast linear plasmid genes have origin close to poxviruses; 2) RNA polymerase subunits, mRNA capping enzyme and putative helicase form a complex *in vivo*, and the putative helicase interacts with plasmid-specific DNA *in vivo*; 3) during transcription initiation 5' polyadenylation of the transcripts occurs *via* RNA polymerase slippage at promoter sequence element, and RNA stem loops influence 3' end formation of the transcripts *in vivo*, probably by the mechanism of intrinsic transcription termination.

**Keywords:** *Kluyveromyces lactis*, linear plasmids, non-canonical RNA polymerase, transcription

# **OBSAH**

<b>SEZNAM POUŽITÝCH ZKRATEK A TERMÍNŮ .....</b>	<b>9</b>
<b>1 ÚVOD .....</b>	<b>11</b>
<b>2 PŘEHLED LITERATURY .....</b>	<b>12</b>
<b>2.1 KVASINKOVÉ CYTOPLASMATICKÉ LINEÁRNÍ PLASMIDY .....</b>	<b>12</b>
2.1.1 Transkripce pGKL plasmidů .....	14
2.1.1.1 RNA polymeráza.....	15
2.1.1.2 mRNA <i>capping</i> enzym .....	16
2.1.1.3 DExD/H box helikáza .....	19
<b>2.2 POXVIRY .....</b>	<b>21</b>
2.2.1 Organizace genomu a nomenklatura genů viru <i>Vaccinia</i> .....	21
2.2.2 Struktura virionu viru <i>Vaccinia</i> .....	23
2.2.3 Životní cyklus viru <i>Vaccinia</i> .....	24
2.2.4 Replikace DNA viru <i>Vaccinia</i> .....	25
2.2.5 Regulace genové exprese a transkripce viru <i>Vaccinia</i> .....	27
2.2.5.1 RNA polymeráza.....	28
2.2.5.2 Posttranskripční modifikace mRNA.....	30
2.2.6 Transkripce časných genů viru <i>Vaccinia</i> .....	32
2.2.6.1 VETF (DExD/H box helikáza) .....	35
2.2.6.2 mRNA <i>capping</i> enzym (VTF) .....	36
2.2.6.3 NPH I (DExD/H box helikáza) .....	38
2.2.6.4 NPH II.....	40
2.2.7 Transkripce středních genů viru <i>Vaccinia</i> .....	41
2.2.7.1 VITF-1, VITF-2, VITF-3, mRNA <i>capping</i> enzym.....	44
2.2.7.2 TATA-box vazebný protein.....	45
2.2.7.3 Negativní elongační faktor A18.....	45
2.2.7.4 Pozitivní elongační faktory G2, J3 a H5.....	47
2.2.8 Transkripce pozdních genů viru <i>Vaccinia</i> .....	49
2.2.8.1 VLTF-1, VLTF-2, VLTF-3, VLTF-4 .....	50
2.2.8.2 VLTF-X, hnRNP A2/RBM3, TATA-box vazebný protein.....	52
<b>3 CÍLE PRÁCE.....</b>	<b>53</b>
<b>4 MATERIÁL A METODY .....</b>	<b>54</b>
<b>4.1 MATERIÁL.....</b>	<b>54</b>
4.1.1 Chemikálie.....	54
4.1.2 Roztoky a pufry .....	57
4.1.2.1 Pufry pro purifikaci proteinů a imunoprecipitaci.....	58
4.1.3 Komerční pufry .....	59
4.1.4 Enzymy a jiné proteiny.....	60
4.1.5 Komerční soupravy .....	60
4.1.6 Standardy molekulových hmotností.....	61
4.1.7 Plasmidy a vektory .....	62
4.1.8 Oligonukleotidové primery.....	64
4.1.9 Protilátky .....	67

4.1.10 Bakteriální kmeny.....	68
4.1.11 Kvasinkové kmeny.....	68
4.1.12 Kultivační média .....	69
4.1.12.1 Bakteriální kultivační média .....	69
4.1.12.2 Kvasinková kultivační média .....	70
<b>4.2 METODY .....</b>	<b>71</b>
4.2.1 Práce s mikroorganismy .....	71
4.2.1.1 Kultivace bakterií .....	71
4.2.1.2 Kultivace kvasinek .....	71
4.2.1.3 Skladování bakteriálních a kvasinkových kmenů .....	71
4.2.1.4 Měření růstu mikroorganismů.....	71
4.2.1.5 Indukce bakteriálních promotorů s <i>lac</i> operátorem.....	72
4.2.2 Práce s DNA.....	72
4.2.2.1 Izolace DNA.....	72
4.2.2.2 Měření koncentrace DNA .....	79
4.2.2.3 Elektroforéza DNA .....	80
4.2.2.4 Modifikace DNA.....	81
4.2.2.5 Vnášení DNA do mikroorganismů.....	84
4.2.2.6 Polymerázová řetězová reakce (PCR) .....	85
4.2.2.7 Polymerázová řetězová reakce v reálném čase (real-time PCR) .....	86
4.2.2.8 Sekvenování DNA .....	87
4.2.3 Práce s RNA .....	89
4.2.3.1 Izolace RNA .....	89
4.2.3.2 Měření koncentrace RNA.....	91
4.2.3.3 Elektroforéza RNA .....	91
4.2.3.4 Modifikace RNA .....	92
4.2.3.5 Transkripční reakce <i>in vitro</i> .....	93
4.2.3.6 Reverzní transkripce .....	94
4.2.4 Práce s proteiny .....	94
4.2.4.1 Příprava bakteriálních lyzátů .....	94
4.2.4.2 Purifikace proteinů s GST-epitopem z bakteriálních a kvasinkových lyzátů .....	95
4.2.4.3 Purifikace proteinů s HIS-epitopem z bakteriálních a kvasinkových lyzátů.....	96
4.2.4.4 Příprava kvasinkových lyzátů .....	97
4.2.4.5 Imunoprecipitace proteinů z kvasinek.....	99
4.2.4.6 Imunoprecipitace proteinů s GFP-epitopem z kvasinkových lyzátů.....	100
4.2.4.7 Imunoprecipitace proteinů s GST-epitopem z kvasinkových lyzátů.....	101
4.2.4.8 Polyakrylamidová proteinová elektroforéza SDS-PAGE.....	101
4.2.4.9 <i>Western blot</i> .....	102
4.2.4.10 Chemiluminiscenční detekce.....	103
4.2.4.11 Barvení polyakrylamidových gelů.....	103
4.2.5 Práce s viru podobnými částicemi .....	104
4.2.5.1 Izolace viru podobných částic.....	104
4.2.6 Práce s obrazovými daty .....	107
4.2.7 Bioinformatické analýzy .....	107
<b>5 VÝSLEDKY .....</b>	<b>109</b>
<b>5.1 KVASINKOVÉ CYTOPLASMATICKÉ LINEÁRNÍ PLASMIDY .....</b>	<b>109</b>
5.1.1 SAM vazebné místo mRNA <i>capping</i> enzymu je nezbytné pro fungování pGKL plasmidů <i>in vivo</i> .....	109

5.1.1.1 Příprava kmene s bodovými mutacemi methyltransferázové domény .....	109
5.1.2 Komponenty transkripčního aparátu pGKL plasmidů interagují <i>in vivo</i> .....	112
5.1.2.1 Příprava kmenů kódujících vybrané proteiny se značkami.....	113
5.1.2.2 Testování proteinových interakcí pomocí koimunoprecipitace a hmotnostní spektrometrie .....	118
5.1.2.3 Testování proteinových interakcí pomocí koimunoprecipitace a imunodetekce .....	123
5.1.2.4 Interakce transkripčních faktorů s RNA polymerázou není závislá na nukleových kyselinách .....	127
5.1.2.5 Předpokládné buněčné proteiny interagující s plasmidovým transkripčním komplexem	128
5.1.3 Produkce RNA polymerázy pGKL plasmidů v bakteriálním expresním systému není úspěšná.....	129
5.1.3.1 Příprava duálních expresních vektorů.....	130
5.1.3.2 Purifikace a detekce rekombinantních proteinů.....	133
5.1.4 Imunoprecipitovaný transkripční komplex pGKL plasmidů vykazuje malou či žádnou RNA polymerázovou aktivitu <i>in vitro</i> .....	134
5.1.5 Imunoprecipitovaná RNA polymeráza a předpokládaná helikáza pGKL plasmidů nevykazuje DNA vazebnou aktivitu <i>in vitro</i> .....	136
5.1.6 Předpokládaná helikáza pGKL plasmidů vykazuje specifickou DNA vazebnou aktivitu <i>in vivo</i> .....	138
5.1.6.1 Kvantifikace chromatinové imunoprecipitace pomocí real-time PCR.....	140
5.1.7 RNA polymeráza pGKL plasmidů má poxvirový původ.....	142
5.1.8 RNA polymeráza pGKL plasmidů má unikátní architekturu .....	144
5.1.9 Promotory pGKL plasmidů mají poxvirový původ .....	147
5.1.10 Promotory pGKL plasmidů obsahují iniciátorový region .....	149
5.1.11 Klouzání RNA polymerázy v iniciátorovém regionu vede k 5' polyadenylaci plasmidových mRNA <i>in vivo</i> .....	151
5.1.12 Transkripty pGKL plasmidů obsahují 5' čepičkovou strukturu.....	155
5.1.13 Mutace v iniciátorovém regionu promotoru vede ke snížené expresi genu <i>in vivo</i> ..	158
5.1.14 Časný promotor viru <i>Vaccinia</i> s iniciátorovým regionem nefunguje jako promotor pGKL plasmidů <i>in vivo</i> .....	160
5.1.15 Definovaná tvorba 3' konců plasmidových mRNA není závislá na promotoru a kódující sekvenci genu <i>in vivo</i> .....	162
5.1.16 3' konce plasmidových mRNA se nachází poblíž předpokládaných RNA vláseň .....	165
5.1.17 RNA vlásenky se podílí na tvorbě 3' konců plasmidových mRNA <i>in vivo</i> .....	167
5.1.18 Transkripty pGKL plasmidů obsahují předpokládaný 3' oligo(U) konec.....	170
5.1.19 Gen <i>K2ORF9</i> pGKL plasmidů kóduje předpokládaný <i>scaffold</i> protein.....	171
5.1.20 Gen <i>D13L</i> viru <i>Vaccinia</i> nedokáže zastoupit funkci genu <i>K2ORF9</i> pGKL plasmidů <i>in vivo</i> .....	173
5.1.21 Předpokládaný <i>scaffold</i> protein <i>K2ORF9p</i> interaguje s transkripčním aparátem pGKL plasmidů <i>in vivo</i> .....	175
5.1.22 Předpokládané viru podobné částice pGKL plasmidů se nepodařilo izolovat ultracentrifugací.....	179
<b>5.2 POXVIRY .....</b>	<b>182</b>
5.2.1 Transkripty viru <i>Vaccinia</i> obsahující 5' poly(A) leader často neobsahují 5' mRNA čepičku .....	182

5.2.2	Pozdní transkripty viru <i>Vaccinia</i> obsahující 5' mRNA čepičku mají kratší 5' poly(A) leader.....	187
5.2.3	3' konce mRNA postreplikativních genů viru <i>Vaccinia</i> se dají mapovat .....	190
5.2.4	3' konce mRNA postreplikativních genů viru <i>Vaccinia</i> se nachází poblíž předpokládaných RNA vlásenek .....	193
5.2.5	Předpokládané RNA vlásenky postreplikativních genů viru <i>Vaccinia</i> jsou schopny indukovat definovanou tvorbu 3' konců mRNA pGKL plasmidů <i>in vivo</i> .....	194
<b>6</b>	<b>DISKUZE .....</b>	<b>197</b>
<b>6.1</b>	<b>KVASINKOVÉ CYTOPLASMATICKÉ LINEÁRNÍ PLASMIDY .....</b>	<b>197</b>
6.1.1	Integrita aktivního místa methyltransferázy mRNA <i>capping</i> enzymu je nezbytná pro udržení pGKL plasmidů <i>in vivo</i> .....	197
6.1.2	Složení transkripčního aparátu pGKL plasmidů <i>in vivo</i> .....	199
6.1.3	Produkce RNA polymerázy pGKL plasmidů v bakteriálním expresním systému....	201
6.1.4	Testování aktivit transkripčního aparátu pGKL plasmidů <i>in vitro</i> a <i>in vivo</i> .....	203
6.1.5	Architektura RNA polymerázy pGKL plasmidů .....	206
6.1.6	Evoluční původ transkripčního aparátu pGKL plasmidů .....	208
6.1.7	Iniciace transkripce a složení 5' konců mRNA pGKL plasmidů .....	210
6.1.8	Terminace transkripce a složení 3' konců mRNA pGKL plasmidů.....	214
6.1.9	Předpokládané viru podobné částice pGKL plasmidů.....	216
6.1.10	Evoluční původ kvasinkových lineárních plasmidů.....	217
6.1.11	Translace mRNA pGKL plasmidů .....	222
<b>6.2</b>	<b>POXVIRY .....</b>	<b>228</b>
6.2.1	Složení 5' konců mRNA viru <i>Vaccinia</i> .....	228
6.2.2	Translace mRNA viru <i>Vaccinia</i> .....	230
6.2.3	Složení 3' konců mRNA a terminace transkripce postreplikativních genů viru <i>Vaccinia</i> .....	233
<b>7</b>	<b>SOUHRN.....</b>	<b>237</b>
<b>8</b>	<b>SEZNAM PUBLIKACÍ .....</b>	<b>238</b>
<b>8.1</b>	<b>PUBLIKACE.....</b>	<b>238</b>
<b>8.2</b>	<b>KONFERENCE.....</b>	<b>238</b>
8.2.1	Přednášky .....	238
8.2.2	Postery .....	238
<b>8.3</b>	<b>PODÍL STUDENTA NA PUBLIKACÍCH.....</b>	<b>239</b>
<b>9</b>	<b>SEZNAM POUŽITÉ LITERATURY .....</b>	<b>240</b>
<b>10</b>	<b>VYBRANÉ PUBLIKACE .....</b>	<b>261</b>
<b>10.1</b>	<b>KVASINKOVÉ CYTOPLASMATICKÉ LINEÁRNÍ PLASMIDY.....</b>	<b>261</b>
10.1.1	Transcription apparatus of the yeast virus-like elements: Architecture, function, and evolutionary origin .....	261
10.1.2	Messenger RNAs of the yeast virus-like elements contain non-templated 5' poly(A) leaders and their expression is independent of eIF4E and Pab1 .....	295

# SEZNAM POUŽITÝCH ZKRATEK A TERMÍNŮ

<b>Å</b>	Ångström ( $10^{-10}$ m)	Ångström ( $10^{-10}$ m)
<b>ACT</b>	actin gene	gen aktinu
<b>ATP</b>	adenosine-5'-triphosphate	adenosin-5'-trifosfát
<b>dATP</b>	2'-deoxyadenosine-5'-triphosphate	2'-deoxyadenosin-5'-trifosfát
<b>bp(s)</b>	base pairs	páry bází
<b>cDNA</b>	complementary DNA	komplementární DNA
<b>CE</b>	promoter core element	element jádra promotoru
<b>Ci</b>	curie ( $3.7 \times 10^{10}$ Bq)	curie ( $3,7 \times 10^{10}$ Bq)
<b>COX1</b>	cytochrom c oxidase subunit 1 gene	gen podjednotky 1 cytochrom c oxidázy
<b>COX2</b>	cytochrom c oxidase subunit 2 gene	gen podjednotky 2 cytochrom c oxidázy
<b>cpm</b>	counts per minutes	počet impulzů za minutu
<b>Ct</b>	cycle of threshold	cyklus prahové hodnoty
<b><i>E. coli</i></b>	<i>Escherichia coli</i>	<i>Escherichia coli</i>
<b>EGFP</b>	enhanced green fluorescent protein	posílený zelený fluorescenční protein
<b>eIF3</b>	eukaryotic translation initiation factor 3	eukaryotický translační iniciační faktor 3
<b>eIF4A</b>	eukaryotic translation initiation factor 4A	eukaryotický translační iniciační faktor 4A
<b>eIF4E</b>	eukaryotic translation initiation factor 4E	eukaryotický translační iniciační faktor 4E
<b>eIF4F</b>	eukaryotic translation initiation factor 4F	eukaryotický translační iniciační faktor 4F
<b>eIF4G1</b>	eukaryotic translation initiation factor 4 gamma 1	eukaryotický translační iniciační faktor 4 gama 1
<b>GFP</b>	green fluorescent protein	zelený fluorescenční protein
<b>HGT1</b>	high-affinity glucose transporter gene	gen vysokoafinitního transportéru glukosy
<b>hnRNP A2</b>	heterogeneous nuclear ribonucleoprotein A2	heterogenní jaderný ribonukleoprotein A2
<b>H-NS</b>	histon-like nucleoid-structuring protein	protein strukturující nukleoid podobný histonu
<b>ChIP</b>	chromatin immunoprecipitation	chromatinová imunoprecipitace
<b>IBT</b>	isatin-β-thiosemicarbazone	isatin-β-thiosemikarbazon
<b>IFO</b>	Institute for Fermentation, Osaka	Institut pro fermentaci, Osaka
<b>INR</b>	initiator region	iniciátorový region
<b>IP</b>	immunoprecipitation	imunoprecipitace
<b><i>K. lactis</i></b>	<i>Kluyveromyces lactis</i>	<i>Kluyveromyces lactis</i>
<b>kDa</b>	kilodalton ( $1.66 \times 10^{-24}$ kg)	kilodalton ( $1,66 \times 10^{-24}$ kg)
<b>mRNA</b>	messenger RNA	mediátorová RNA
<b>NPH I</b>	nucleosidetriphosphate phosphohydrolase I	nukleosidtrifosfátfosfohydroláza I
<b>NPH II</b>	nucleosidetriphosphate phosphohydrolase II	nukleosidtrifosfátfosfohydroláza II
<b>nt</b>	nucleotide	nukleotid
<b>NTP</b>	ribonucleoside-5'-triphosphate	ribonukleosid-5'-trifosfát
<b>OD</b>	optical density	optická denzita
<b>ORF</b>	open reading frame	otevřený čtecí rámec
<b>PABP1</b>	poly(A) binding protein 1	poly(A) vazebný protein 1
<b>PAP</b>	poly(A) polymerase	poly(A) polymeráza
<b>PAGE</b>	polyacrylamide gel electrophoresis	polyakrylamidová gelová elektroforéza
<b>PCR</b>	polymerase chain reaction	polymerázová řetězová reakce
<b>PFU</b>	plaque forming unit	jednotka tvořící plak
<b>POLADS</b>	polyadenylated RNA sequences	polyadenylované RNA sekvence
<b>RACE</b>	rapid amplification of cDNA ends	rychlá amplifikace konců cDNA
<b>RACK1</b>	receptor for activated C kinase 1	receptor pro aktivovanou kinázu C 1
<b>RAP94</b>	RNA polymerase-associated protein 94 kDa	94 kDa protein asociovaný s RNA polymerázou
<b>RBM3</b>	RNA-binding motif protein 3	protein s RNA vazebným motivem 3
<b>RIM1</b>	Replication in mitochondria 1	replikace v mitochondriích 1
<b>RLM-RACE</b>	RNA ligase mediated RACE	RACE zprostředkované RNA ligázou
<b>RNAP</b>	RNA polymerase	RNA polymeráza



<b><i>S. cerevisiae</i></b>	<i>Saccharomyces cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
<b>SAM</b>	S-adenosyl-L-methionine	S-adenosyl-L-methionin
<b>SSB</b>	single strand binding	vázající jedno vlákno
<b>ssDNA</b>	single-stranded DNA	jednovláknová DNA
<b>TBP</b>	TATA-box binding protein	TATA-box vazebný protein
<b>TRF1</b>	terminal recognition factor 1	faktor 1 rozeznávající konec
<b>TSS</b>	transcription start site	místo startu transkripce
<b>UCE</b>	upstream control element	kontrolní element proti směru transkripce
<b>UCR</b>	region containing UCS	region obsahující UCS
<b>UCS</b>	upstream conserved sequence	konzervovaná sekvence proti směru transkripce
<b>UTR</b>	untranslated region	nepřekládaná oblast
<b>UV</b>	ultraviolet	ultrafialové záření
<b>VACV</b>	Vaccinia virus	virus <i>Vaccinia</i>
<b>VETF</b>	Vaccinia virus early transcription factor	časný transkripční faktor viru <i>Vaccinia</i>
<b>VITF</b>	Vaccinia virus intermediate transcription factor	střední transkripční faktor viru <i>Vaccinia</i>
<b>VLP</b>	virus-like particles	viru podobné částice
<b>VLTF</b>	Vaccinia virus late transcription factor	pozdní transkripční faktor viru <i>Vaccinia</i>
<b>VTF</b>	Vaccinia virus transcription termination factor	transkripční terminační faktor viru <i>Vaccinia</i>
<b>yEGFP3</b>	yeast enhanced green fluorescent protein 3	kvasinkový posílený zelený fluorescenční protein 3
<b>YY1</b>	Yin Yang 1	Yin Yang 1
<b>ρ<sup>0</sup></b>		absence mitochondriální DNA

## Použité zkratky aminokyselin / aminokyselinových zbytků

<b>A</b>	alanin	<b>I</b>	isoleucin	<b>R</b>	arginin
<b>C</b>	cystein	<b>K</b>	lysin	<b>S</b>	serin
<b>D</b>	aspartát	<b>L</b>	leucin	<b>T</b>	threonin
<b>E</b>	glutamát	<b>M</b>	methionin	<b>V</b>	valin
<b>F</b>	fenylalanin	<b>N</b>	asparagin	<b>W</b>	tryptofan
<b>G</b>	glycin	<b>P</b>	prolin	<b>Y</b>	tyrozin
<b>H</b>	histidin	<b>Q</b>	glutamin	<b>x</b>	jakákoliv

## Použité zkratky nukleosidů / nukleotidů

<b>A</b>	adenosin	<b>B</b>	C nebo G nebo T	<b>C</b>	cytidin
<b>G</b>	guanosin	<b>U</b>	uridin	<b>T</b>	thymidin
<b>N</b>	jakýkoliv				

Všeobecně známé zkratky (například jednotky SI a jednotky od nich odvozené) zde nejsou vypsány. Zkratky použité v textu pouze jednou u obrázků zde také nejsou vypsány. Použité zkratky některých chemikálií a laboratorních enzymů jsou uvedeny v jiné kapitole této práce.

# 1 ÚVOD

Tato disertační práce se zabývá tematikou základního výzkumu lineárních DNA plasmidů v mléčném průmyslu využívané kvasinky *Kluyveromyces lactis*. Některé kmeny této kvasinky obsahují dva plasmidy, nazvané pGKL1 a pGKL2, které jsou pozoruhodné svojí cytoplasmatickou lokalizací a uspořádáním genomu, čímž připomínají cytoplasmatické viry čeledi *Poxviridae* a *Asfarviridae*. Přestože byly pGKL plasmidy popsány již před 37 lety, stále zůstává mnoho nejasností ohledně jejich biologie. To bezesporu souvisí také s tím, že výzkum molekulárních aspektů fungování pGKL plasmidů byl moderní pouze v 80. a 90. letech 20. století, tedy v předsekvenční době, kdy ještě nepřipadalo v úvahu čtení celých genomů buněčných organismů. Výzkum té doby se tedy hojněji zaměřoval na malé kompaktní virové a plasmidové genomy, které bylo snadné sekvenčně charakterizovat. Po konci 90. let 20. století se nové informace o lineárních plasmidech objevují jen ojediněle, což souvisí s tím, že naprostá většina vědeckých týmů, zabývajících se touto tematikou, přešla na jiná, aktuálnější a financovatelná témata. Co ovšem stále zůstává v relativním centru zájmu, je studium mechanismu účinku plasmidy kódovaného *killer* toxinu, který je v současné době velmi podrobně charakterizován.

V kontextu katedry genetiky a mikrobiologie Přírodovědecké fakulty Univerzity Karlovy má studium lineárních plasmidů kvasinek dlouhou tradici; Laboratoř biochemie RNA se tomuto tématu věnuje od svého vzniku se snižující se mírou až dodnes. Téma zapadá do celkové koncepce laboratoře, kterou je výzkum molekulárních mechanismů procesu iniciace translace buněčných a virových mediátorových RNA (mRNA) v eukaryotních buňkách. O mechanismu translace genů kódovaných lineárními plasmidy nebylo nic známo, a to také díky chybějícím informacím o možných posttranskripčních úpravách plasmidových mRNA. Pomocí bioinformatických analýz a pilotních experimentů bylo zjištěno, že pGKL plasmidy kódují všechny nezbytné součásti svého vlastního, pravděpodobně soběstačného, transkripčního aparátu, čímž se podobají výše zmíněným cytoplasmatickým virům. Nicméně molekulární detaily transkripce a tedy i translace nebyly známy. Právě proto se tato práce zaměřila na detailnější charakterizaci transkriptů a komponent transkripčního aparátu lineárních plasmidů, čímž alespoň malým dílem přispěla k osvětlení mechanismu iniciace a terminace transkripce plasmidových genů a také k objasnění evolučního původu plasmidového transkripčního aparátu. V návaznosti na zjištění předpokládaného evolučního původu transkripčního systému kvasinkových lineárních plasmidů se tato práce ve své druhé části zabývá detailnější charakterizací nepřekládaných oblastí transkriptů modelového viru čeledi *Poxviridae*, viru *Vaccinia*. Hlavně díky menšímu množství dostupných publikací, týkajících se transkripčního aparátu lineárních plasmidů, a díky zřejmé podobnosti s transkripčním aparátem virů čeledi *Poxviridae* je literární přehled z velké části zaměřen právě na tento evolučně příbuzný virový transkripční aparát.

## 2 PŘEHLED LITERATURY

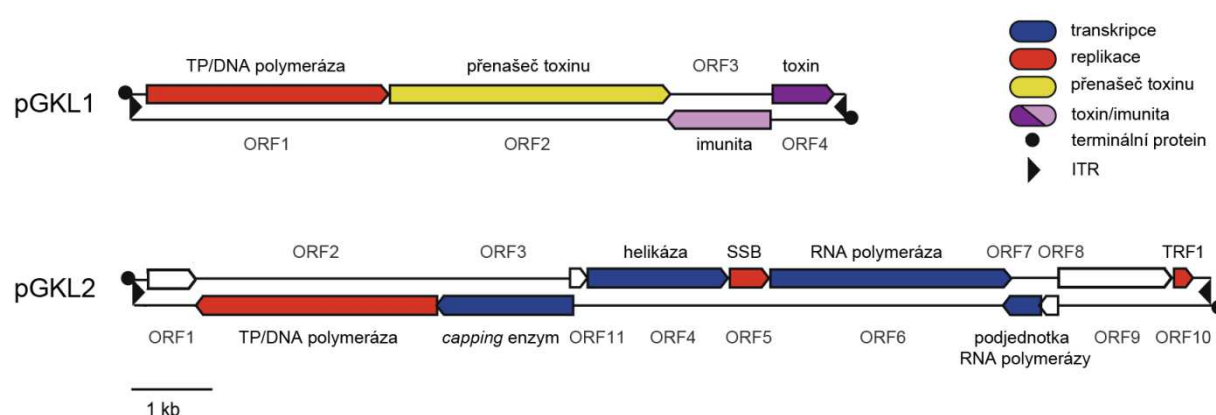
### 2.1 KVASINKOVÉ CYTOPLASMATICKÉ LINEÁRNÍ PLASMIDY

Lineární cytoplasmatické plasmidy byly nalezeny u několika druhů převážně askomycetních kvasinek, patřících do devíti rodů (KLASSEN a MEINHARDT 2007). Při systematickém prohledávání 1805 kvasinkových kmenů byly lineární plasmidy detekovány s frekvencí 1,8 %, což značí, že alespoň u některých druhů jsou poměrně rozšířeny (FUKUHARA 1995). Genetická organizace kvasinkových dvouvláknových DNA plasmidů je poměrně uniformní bez ohledu na taxonomický a geografický původ. Tyto plasmidy často tvoří soustavu dvou nebo tří různých velkých elementů v kvasinkové hostitelské buňce (JESKE *et al.* 2007). Charakteristickými rysy těchto plasmidů jsou terminální proteiny kovalentně spojené s 5' konci jejich DNA, terminální invertované repetice a cytoplasmatická lokalizace (GUNGE *et al.* 1982; KIKUCHI *et al.* 1984; STAM *et al.* 1986). První lineární plasmidy popsány u kvasinky *Kluyveromyces lactis* (*K. lactis*), nazvané pGKL1 (synonymum K1) a pGKL2 (synonymum K2), se staly modelovým systémem pro studium těchto elementů.

Plasmidy *K. lactis* mají kompaktní genom s občasnými překryvy otevřených čtecích rámců (ORF) a velmi vysoký (~74 %) obsah AT bází (SOR a FUKUHARA 1985; TOMMASINO *et al.* 1988). Přítomnost pGKL plasmidů u několika kmenů *K. lactis* je spojena s kvasinkovým *killer* fenotypem (GUNGE *et al.* 1981), který byl rozsáhle studován. Plasmid pGKL1 kóduje heterotrimerní sekretovaný *killer* toxin, nazývaný zymocin, a imunitní složku, chránící hostitelské buňky proti účinkům toxinu (STARK a BOYD 1986; TOKUNAGA *et al.* 1987). Systém toxin/antitoxin přispívá k stabilní dědičnosti obou plasmidů autoselekcí, k jejímuž udržování napomáhá i velmi vysoký obsah AT bází plasmidových otevřených čtecích rámců, díky kterému nedochází k funkční expresi těchto ORF po jejich přenosu do jádra hostitelské buňky (KAST *et al.* 2015). Plasmid pGKL1 je neautonomní a pro jeho udržení v buňkách je nezbytná přítomnost většího autonomního plasmidu pGKL2, který se může v buňkách laboratorně připravených kmenů vyskytovat i samostatně. Plasmid pGKL2 kóduje proteiny s předpokládanými funkcemi v replikaci a transkripci obou plasmidů (STARK *et al.* 1990). Pouze dva otevřené čtecí rámce (ORF2-pGKL1 a ORF1-pGKL2) jsou neesenciální pro udržení obou plasmidů v buňce (SOR a FUKUHARA 1985; SCHAFFRATH *et al.* 1992). Kromě plasmidu pGKL2 jsou známy ještě kompletní nukleotidové sekvence autonomních elementů pSKL *Saccharomyces kluyveri*, pPE1B *Pichia etchellsii*, pPac1-1 *Pichia acaciae* a pKP *Komagataella phaffii* (HISHINUMA a HIRAI 1991; KLASSEN *et al.* 2001; JESKE a MEINHARDT 2006; LOVE *et al.* 2016). Díky obdobnému genetickému uspořádání a velmi vysoké sekvenční identitě jednotlivých genů mohou být

výsledky získané na modelovém systému v *K. lactis* generalizovány. Funkce proteinových produktů většiny otevřených čtecích rámců pGKL plasmidů byla předpovězena pomocí bioinformatických přístupů, avšak pouze některé proteiny byly experimentálně charakterizovány. Orientace a délka otevřených čtecích rámců obou plasmidů je znázorněna na Obrázku 1. Udává se, že se oba plasmidy vyskytují přibližně v 50–100 kopiích na haploidní buňku (GUNGE 1983). Oba plasmidy nesou otevřený čtecí rámec kódující terminální protein a DNA polymerázu, nezbytný pro svoji selektivní replikaci (KIKUCHI *et al.* 1985; KITADA a GUNGE 1988). Terminální protein funguje jako primer pro replikaci a zůstává kovalentně spojen s 5' konci DNA. U kvasinkových lineárních plasmidů je terminální protein kódován N-koncovou částí genu pro DNA polymerázu (TAKEDA *et al.* 1996) a není známo, kdy a jakým způsobem je tento fúzní protein rozdělen. Předpokládá se, že mechanismus replikace pGKL plasmidů je nezávislý na hostitelských proteinech a je obdobný replikaci bakteriofága  $\varphi 29$  z *Bacillus subtilis* nebo virů z rodiny *Adenoviridae* (STARK *et al.* 1990). V předpokládaném modelu iniciace replikace tvoří fúzní protein DNA polymerázy komplex s *terminal recognition factor 1* (TRF1) proteinem, jenž se váže na invertované terminální repetice obou konců plasmidu (MCNEEL a TAMANOI 1991). Současně dojde k vytěsnění původního vlákna, které je stabilizováno vazbou *single stranded binding* (SSB) proteinu (SCHAFFRATH a MEACOCK 2001).

Původ lineárních plasmidů kvasinek zůstává nejasný. Má se za to, že mohou pocházet z endosymbiotické bakterie nebo viru (KEMPKEN *et al.* 1992). Zaznamenaná sekvenční podobnost mezi virovými enzymy a předpokládanými produkty několika otevřených čtecích rámců lineárních plasmidů s očekávanými funkcemi v replikaci a transkripci vedla k tomu, že se dnes kvasinkové lineární plasmidy nazývají virům podobné elementy (SATWIKI *et al.* 2012).



**Obrázek 1:** Schématické znázornění struktury a kódující kapacity lineárních plasmidů kvasinky *Kluyveromyces lactis*. Šipky indikují otevřené čtecí rámce a jejich transkripční směr. Barvy odpovídají uvedeným funkčním kategoriím. Bezbarvé šipky značí otevřené čtecí rámce s neznámou funkcí. Známé nebo předpokládané funkce predikovaných proteinů jsou uvedeny. Uvedené měřítko je tisíc párů bází. TP: terminální protein, ITR: invertované terminální repetice, SSB: *single strand binding* protein, TRF1: *terminal recognition factor 1*. Upraveno podle (KLASSEN a MEINHARDT 2007).

### 2.1.1 Transkripce pGKL plasmidů

Bylo prokázáno, že transkripce plasmidových genů pravděpodobně využívá specifickou RNA polymerázu, neboť je nezávislá jak na mitochondriální RNA polymeráze (GUNGE a YAMANE 1984) tak i na jaderných RNA polymerázách (ROMANOS a BOYD 1988; KÄMPER *et al.* 1989a; STARK *et al.* 1990; KÄMPER *et al.* 1991). Díky experimentům s bakteriálními reportérovými a kvasinkovými jadernými geny, fúzovanými se sekvencemi pGKL plasmidů, byla identifikována konzervovaná sekvence proti směru transkripce (5'-ATNTGA-3') vyskytující se před každým z otevřených čtecích rámců. Konzervovaná sekvence proti směru transkripce (UCS) je nezbytná pro cytoplasmatickou transkripci genu umístěného po směru transkripce (SCHRÜNDER a MEINHARDT 1995; SCHICKEL *et al.* 1996; SCHRÜNDER *et al.* 1996). Tato sekvence je umístěna ve vzdálenosti průměrně 30 nukleotidů před start kodónem, i když se její vzdálenost u jednotlivých ORF může mírně lišit (18–207 nukleotidů) (HISHINUMA a HIRAI 1991; JESKE a MEINHARDT 2006). Bylo prokázáno, že sekvence proti směru transkripce lokalizované distálně od UCS elementu nemají žádný vliv na transkripci, což vylučuje přítomnost dalších promotorových motivů (SCHICKEL *et al.* 1996). UCS element je vysoce konzervovaný mezi všemi kvasinkovými lineárními plasmidy (STARK *et al.* 1990; HISHINUMA a HIRAI 1991; FUKUDA *et al.* 1997; KLASSEN *et al.* 2002; KLASSEN a MEINHARDT 2003; JESKE a MEINHARDT 2006). Jelikož sekvence UCS odvozená z ORF6 plasmidu pPE1B kvasinky *Pichia etchellsii* působí jako plně funkční promotor pro expresi reportérového genu umístěného do plasmidu pGKL1 (KLASSEN *et al.* 2001), je zřejmé, že je UCS sekvence univerzální *cis* regulační složkou specifického transkripčního systému lineárních plasmidů, která je nějakým zatím neznámým mechanismem nezbytná pro iniciaci transkripce, ale její původ a přesná funkce jsou neznámé.

Ačkoliv je exprese genů lineárních plasmidů poměrně slabá, byla určena relativní transkripční aktivita UCS dvou ORF plasmidu pGKL1 a sedmi ORF plasmidu pGKL2 tím, že se sekvence od UCS po iniciační kodón použila k expresi glukosadehydrogenázy, jejíž enzymatická aktivita byla měřena (SCHICKEL *et al.* 1996; LARSEN a MEINHARDT 2000). Nejvyšší transkripční aktivita byla detekována u elementu řídícího transkripci předpokládané velké podjednotky RNA polymerázy a nejnižší aktivita byla detekována u elementu předpokládané DNA polymerázy plasmidu pGKL2, což pravděpodobně koreluje s předpokládanými požadavky systému lineárních plasmidů na množství příslušných proteinů (SCHICKEL *et al.* 1996). Nicméně celková aktivita glukosadehydrogenázy byla ve všech případech nižší než při její expresi v klasických bakteriálních a kvasinkových expresních systémech (HEILMANN *et al.* 1988), takže nízká míra exprese plasmidových genů se mohla vyvinout tak, aby stabilní propagace těchto elementů neposkytovala výraznou zátěž pro hostitelské buňky (SCHICKEL *et al.* 1996).

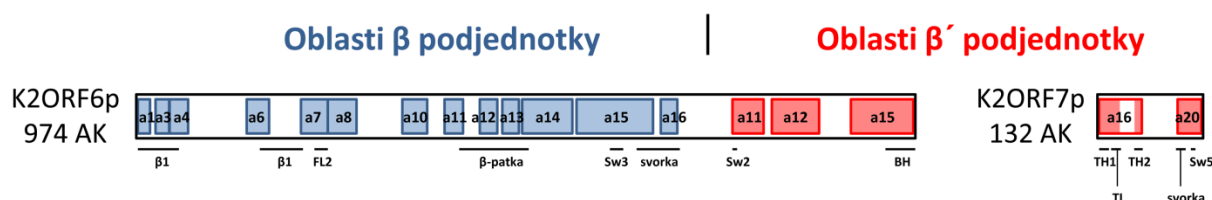
Analýzy iniciace transkripce, provedené na transkriptech všech ORF plasmidu pGKL1 a dvou ORF plasmidu pGKL2 metodou prodlužování primeru a/nebo mapováním S1 endonukleázou, naznačují více iniciačních míst, která se většinou nacházejí ve vzdálenosti 8–16 nukleotidů od UCS sekvence po směru transkripce, kdy jedno z těchto míst je většinou preferováno (SOR a FUKUHARA 1985; ROMANOS a BOYD 1988; STARK *et al.* 1990; SCHAFFRATH a MEACOCK 1995; JESKE *et al.* 2006). Pomocí metody *northern blot* byly detekovány transkripty odvozené od všech ORF plasmidu pGKL1 a šesti ORF plasmidu pGKL2, jejichž zjištěná délka odpovídá pouze monocistronním transkriptům, což naznačuje existenci definovaného, leč dosud neznámého mechanismu terminace transkripce (ROMANOS a BOYD 1988; TOMMASINO *et al.* 1988; SCHAFFRATH *et al.* 1995b; SCHAFFRATH *et al.* 1997; JESKE *et al.* 2006). Mapování 3' konců transkriptů všech ORF plasmidu pGKL1 S1 endonukleázou nebylo příliš úspěšné, protože signály byly široce rozprostřené, což značí, že 3' nepřekládané oblasti jsou heterogenní v délce mezi jednotlivými transkripty (SOR a FUKUHARA 1985). Díky délce a vzájemné překryvnosti detekovaných 3' konců transkriptů ORF2 a ORF3 plasmidu pGKL1 bylo navrženo, že transkripční terminátory jsou jednosměrné (STARK *et al.* 1990). Dodnes však zůstává mechanismus terminace neznámý, stejně jako potenciální 3' koncové mRNA modifikace, jako je polyadenylace. Hlavní komponenty transkripčního aparátu lineárních plasmidů jsou unikátní RNA polymeráza (K2ORF6p a K2ORF7p), mRNA *capping* enzym (K2ORF3p) a pravděpodobná DEAD/H box helikáza (K2ORF4p).

### 2.1.1.1 RNA polymeráza

Produkty otevřených čtecích rámců 6 a 7 plasmidu pGKL2 pravděpodobně tvoří nekanonickou RNA polymerázu (RNAP). V otevřeném čtecím rámci 6 byla nalezena sekvenční podobnost s třemi oblastmi dvou největších podjednotek ( $\beta$  a  $\beta'$  u bakterií) kanonických vícepodjednotkových RNA polymeráz (WILSON a MEACOCK 1988). Podobnost K2ORF6p k  $\beta$  a  $\beta'$  podjednotkám byla rozšířena na 12 ze 36 konzervovaných oblastí, které sdílejí všechny bakteriální, archeální i eukaryotické RNAP (LANE a DARST 2010a; RUPRICH-ROBERT a THURIAUX 2010), a unikátní architektura tohoto enzymu se tak podobá fúzi  $\beta$  podjednotky s částí  $\beta'$  podjednotky (Obrázek 2). Sekvenční podobnost K2ORF6p k  $\beta$  podjednotce se mi podařilo v rámci diplomové práce rozšířit o další 4 všeobecně konzervované oblasti RNAP (SÝKORA 2013). Fúze  $\beta$  a  $\beta'$  podjednotky RNAP je běžná u některých parazitických bakterií, například u rodu *Helicobacter* a *Wolbachia* (LANE a DARST 2010a). U proteinu K2ORF6p však chybí zejména první polovina konzervovaných oblastí  $\beta'$  podjednotky (RUPRICH-ROBERT a THURIAUX 2010). Kromě zjištění, že se jedná o nepostradatelný gen pro udržení pGKL



plasmidů v buňkách (SCHAFFRATH *et al.* 1995b), nebyl gen *ORF6* ani jeho proteinový produkt nikterak charakterizován. V otevřeném čtecím rámci 7 plasmidu pGKL2 byla nalezena sekvenční podobnost ke 2 konzervovaným oblastem  $\beta'$  podjednotky, které se u většiny kanonických RNA polymeráz nacházejí na jejich C-konci (SCHAFFRATH *et al.* 1997). Rozdělení konzervovaných oblastí  $\beta'$  podjednotky na dva polypeptidy ve stejném místě, jako je tomu u K2ORF6p a K2ORF7p, je známo u kanonických RNA polymeráz většiny druhů *Archaea* (KWAPISZ *et al.* 2008), přičemž rozdělení nemá na funkci  $\beta'$  podjednotky vliv (SEVERINOV *et al.* 1996). Gen *ORF7* plasmidu pGKL2 je taktéž nepostradatelný pro udržení pGKL plasmidů v buňkách (SCHAFFRATH *et al.* 1997). Navíc byl K2ORF7p nadprodukovan v bakteriálním expresním systému se značkou c-Myc na svém C-konci a protilátkou proti K2ORF7p-c-Myc fúznímu proteinu byl metodou *western blot* identifikován K2ORF7p v lyzátu kvasinkových buněk kmene nesoucího pGKL plasmidy (SCHAFFRATH *et al.* 1997). K další charakterizaci tohoto proteinu nedošlo. Předpokládá se, že geny *ORF6* a *ORF7* plasmidu pGKL2 by mohly být odvozeny z ancestrálních kvasinkových genů kódujících orthology  $\beta$  a  $\beta'$  podjednotky eukaryotické RNAP I nebo II či III (RUPRICH-ROBERT a THURIAUX 2010). Další teorie říká, že K2ORF6p by mohl být zástupce dávné diverzifikace vícepodjednotkových RNAP z dob, kdy byla  $\beta$  a  $\beta'$  podjednotka součástí jednoho proteinu (IYER a ARAVIND 2012). Geny pro nekanonickou RNAP se u všech dosud charakterizovaných kvasinkových lineárních plasmidů vyskytují na autonomním plasmidu, což je jedním z důvodů nesamostatnosti menších neautonomních plasmidů (JESKE *et al.* 2007).



**Obrázek 2:** Schématické zobrazení oblastí RNA polymerázy lineárních plasmidů kvasinek vykazujících podobnost ke konzervovaným oblastem katalytických podjednotek kanonických vícepodjednotkových RNAP. Konzervované oblasti  $\beta$  podjednotky (modře) a  $\beta'$  podjednotky (červeně) přítomné v K2ORF6p a K2ORF7p jsou zobrazeny v odpovídajícím měřítku a pojmenovány podle (LANE a DARST 2010a). Umístění důležitých strukturních prvků vícepodjednotkových RNAP je vyznačeno pod schématem. BH: mostový helix, FL2: vidličková smyčka 2, Sw2: přepínač 2, Sw3: přepínač 3, Sw5: přepínač 5, TH1: spouštěcí helix 1, TH2: spouštěcí helix 2, TL: spouštěcí smyčka. Upraveno podle (SÝKORA 2013).

### 2.1.1.2 mRNA *capping* enzym

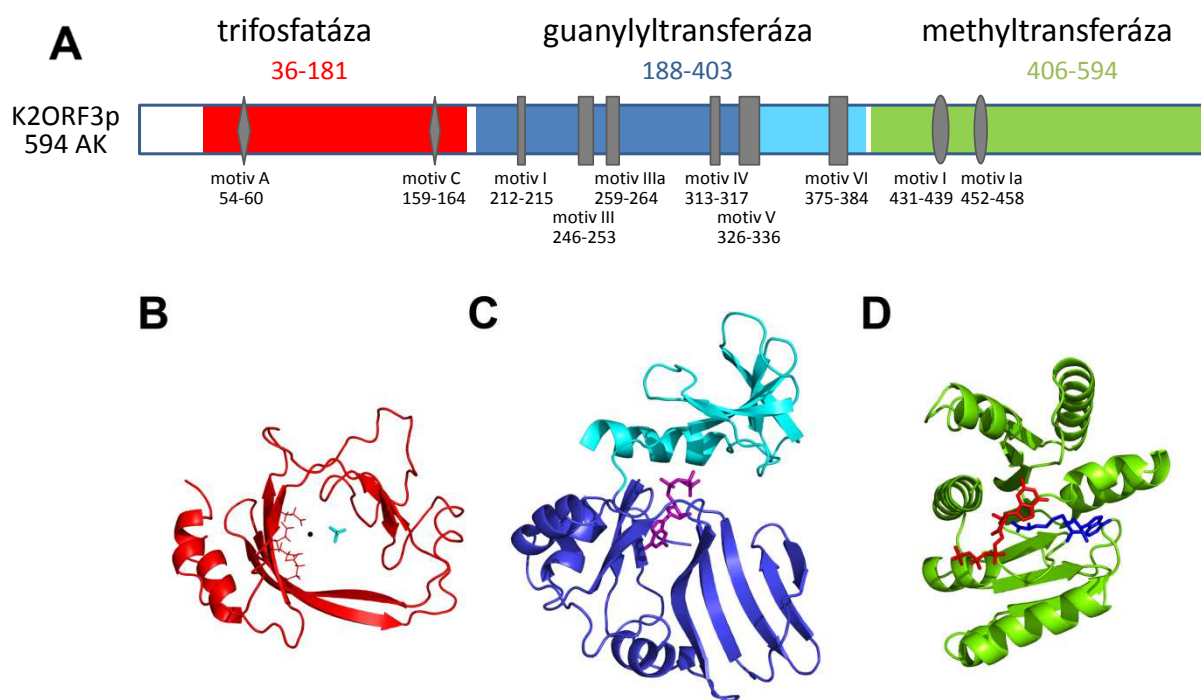
Sekvence otevřeného čtecího rámce 3 plasmidu pGKL2 vykazuje nápadnou sekvenční podobnost k virovým tripartitním mRNA *capping* enzymům, zajišťujícím syntézu N<sup>7</sup>-methylguanosinové čepičkové struktury mRNA viz Obrázek 3 (LARSEN *et al.* 1998).

Přítomnost kovalentně vázané čepičkové struktury na 5' koncích mediátorové RNA je běžným znakem eukaryotických mRNA a také RNA mnoha eukaryotických virů. Pro tvorbu základní N<sup>7</sup>-methylguanosinové čepičkové struktury (označované jako *cap* 0) jsou vyžadovány tři enzymatické aktivity, přičemž existují značné rozdíly v genetické organizaci, struktuře a mechanismu katalýzy mezi buněčnými a virovými mRNA *capping* enzymy, pro přehled viz (GHOSH a LIMA 2010). K2ORF3p vykazuje sekvenční podobnost s mRNA *capping* enzymem viru *Vaccinia*, ve kterém jsou domény zodpovědné za RNA-trifosfátázovou, RNA-guanylyltransferázovou a RNA guanin-N<sup>7</sup>-methyltransferázovou enzymatickou aktivitu, nutnou pro tvorbu čepičkové struktury, přítomné na jednom polypeptidu kódovaném genem *D1R* (SHUMAN 1995; LARSEN *et al.* 1998). Stejně tripartitní uspořádání domén jako u proteinu D1 mají mRNA *capping* enzymy i některých dalších nukleocytoplasmatických DNA virů z čeledi *Mimiviridae*, *Marseilleviridae* a *Asfarviridae* (BENARROCH *et al.* 2008; BOYER *et al.* 2009; RODRÍGUEZ a SALAS 2013). Nejvíce studovaným tripartitním mRNA *capping* enzymem je ovšem právě protein D1 z poxviru *Vaccinia*, který však vykazuje pouze velmi slabou RNA guanin-N<sup>7</sup>-methyltransferázovou aktivitu, jež je alostericky stimulována (více než 30násobně) heterodimerizací s menším proteinem, kódovaným genem *D12L* (HIGMAN *et al.* 1994; MAO a SHUMAN 1994; SCHWER *et al.* 2006). Komplex D1 a D12 proteinu u viru *Vaccinia* je někdy také nazýván jako virový terminační faktor (VTF), neboť kromě role v tvorbě čepičkové struktury na 5' koncích virové mRNA funguje jako terminační faktor transkripce časných genů a také jako iniciační faktor transkripce středních genů (HARRIS *et al.* 1993; LUO *et al.* 1995).

Gen *K2ORF3* je nepostradatelný pro udržení pGKL plasmidů v buňkách (LARSEN *et al.* 1998). V N-koncové části K2ORF3p byly nalezeny sekvenční motivy charakteristické pro RNA-trifosfátázy s tunelovou strukturou, závislé na dvoumocných kationtech, které se vyskytují u nižších eukaryot a některých virů (LARSEN *et al.* 1998; LIMA *et al.* 1999; HAUSMANN *et al.* 2002). Střední část K2ORF3p obsahuje KxDG motiv a 5 dalších konzervovaných motivů, charakteristických pro rodinu mononukleotidyltransferáz (SHUMAN a SCHWER 1995; LARSEN *et al.* 1998). V C-koncové části K2ORF3p byly nalezeny 2 motivy s velkou sekvenční podobností k podrodině Abd1 methyltransferáz, včetně potenciálního vazebného místa motivu I pro S-adenosyl-L-methionin (SAM), jenž slouží jako donor methylové skupiny pro tento druh methylování (WANG a SHUMAN 1997; LARSEN *et al.* 1998). Experimentálně již byla potvrzena RNA-trifosfátázová a RNA-guanylyltransferázová aktivita K2ORF3p *in vitro*, stejně jako esencialita aminokyselinového zbytku K<sup>212</sup> motivu I (KxDG) guanylyltransferázové domény pro kovalentní vazbu GMP *in vitro* (TIGGEMANN *et al.* 2001). Methyltransferázovou aktivitu produktu K2ORF3p se navzdory několika nezávislým pokusům nepodařilo *in vitro* prokázat (TIGGEMANN *et al.* 2001; VOPÁLENSKÝ 2007). Stále není známo, zda je potřebná případná druhá podjednotka



*capping* enzymu pro účinnou methyltransferázovou aktivitu K2ORF3p, jako tomu je u viru *Vaccinia*. Vhodnými kandidáty tohoto potenciálního alosterického aktivátoru by mohly být otevřené čtecí rámce pGKL2 se zatím neznámými funkcemi. Na základě fylogenetické analýzy aminokyselinových sekvencí guanylyltransferázových domén se jeví geny ORF3 kvasinkových lineárních plasmidů nejpodobnější právě genům *D1R* u *Poxviridae* (JESKE *et al.* 2007).

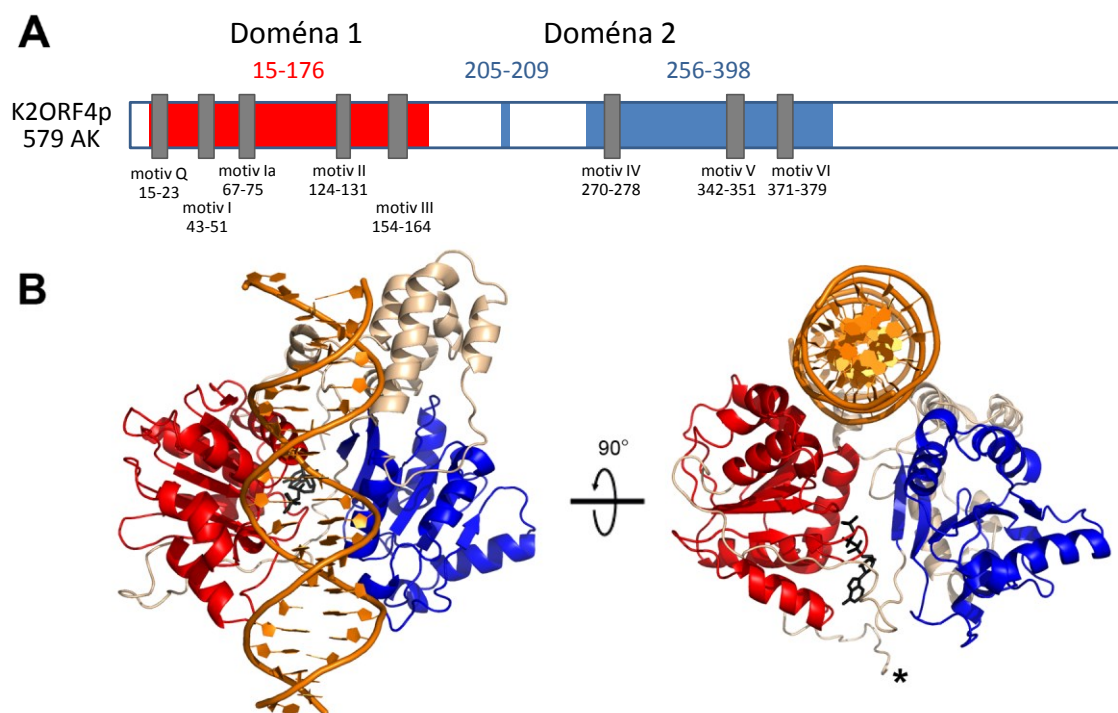


**Obrázek 3:** Schématické zobrazení sekvence a predikované terciární struktury mRNA *capping* enzymu lineárních plasmidů kvasinek vykazující podobnost k virovým tripartitním *capping* enzymům. (A) 5' RNA-trifosfatázová doména (červeně), mRNA-guanylyltransferázová doména (modře a tyrkysově) a mRNA *cap* (guanin-N<sup>7</sup>)-methyltransferázová doména (zeleně) K2ORF3p je zobrazena v odpovídajícím měřítku. Konzervované motivy trifosfatázy (šedé kosočtverce) jsou označeny a pojmenovány podle (LIMA *et al.* 1999; PEI *et al.* 1999). Konzervované motivy guanylyltransferázy (šedé obdélníky) jsou popsány podle (LARSEN *et al.* 1998). Konzervované motivy I a Ia methyltransferázy (šedé ovály) jsou popsány podle (WANG a SHUMAN 1997). (B) Model terciární struktury K2ORF3p trifosfatázy (červeně), pokrývající aminokyselinové zbytky 36–181, s Mn<sup>2+</sup> iontem (černě) a analogem γ-fosfátu (tyrkysově). Postranní řetězce čtyř univerzálně konzervovaných kyselých aminokyselinových zbytků motivu A a C, interagujících s dvoumocným kationtem, jsou zobrazeny prostorovým rozložením atomů. Model byl vytvořen homologním modelováním serverem I-TASSER podle několika templátů. Zobrazení ligandů je založeno na struktuře RNA-trifosfatázy Cet1 v komplexu s Mn<sup>2+</sup> a sulfátem (PDB: 1D8H), s níž byl model strukturně porovnán pomocí DaliLite v. 3 a vizualizován pomocí PyMOL 1.3. (C) Model terciární struktury K2ORF3p guanylyltransferázy obsahující nukleotidyltransferázovou doménu (modře) a oligonukleotid vázající doménu (tyrkysově), pokrývající aminokyselinové zbytky 188–403, s GTP (fialově). Model byl vytvořen homologním modelováním serverem Robetta podle RNA-guanylyltransferázy Cgt1 (PDB: 1P16). Zobrazení GTP je založeno na struktuře RNA-guanylyltransferázy viru PBCV-1 v komplexu s GTP (PDB: 1CKM). (D) Model terciární struktury K2ORF3p methyltransferázy (zeleně), pokrývající aminokyselinové zbytky 406–594, s 7-methyl-guanosin 5' trifosfátem (červěně) a S-adenosyl-L-homocysteinem (modře). Model byl vytvořen homologním modelováním serverem Robetta podle mRNA *cap* (guanin-N<sup>7</sup>)-methyltransferázy Ecm1 (PDB: 1RI5). Struktura ligandů pochází ze struktury Ecm1 (PDB ID: 1RI1).

### 2.1.1.3 DExD/H box helikáza

Sekvence otevřeného čtecího rámce 4 plasmidu pGKL2 vykazuje nápadnou sekvenční podobnost k virovým helikázám superrodiny 2 z rodiny DExD/H box helikáz podílejících se na transkripci (BAYLIS *et al.* 1993). Tento otevřený čtecí rámec je nepostradatelným genem pro udržení pGKL plasmidů v buňkách (SCHAFFRATH *et al.* 1995a). Helikázy nukleových kyselin se dělí na 6 superrodin, přičemž helikázy superrodiny 1 a 2 jsou monomerické proteiny sdílející 7 hlavních sekvenčně konzervovaných motivů a charakteristické dvoulaločnaté jádro enzymu tvořené dvěma doménami (KADARÉ a HAENNI 1997; LEITÃO *et al.* 2015). Helikázy rodiny DExD/H box obsahují ještě konzervovaný motiv Q, skládající se z několika aminokyselinových zbytků (TANNER *et al.* 2003). Helikázová superrodina 2 je největší a nejvíce diverzifikovaná skupina, jejíž zástupci se podílí na procesech transkripce, reparace DNA, remodelace chromatinu a metabolismu RNA (BYRD a RANEY 2012). Všechny enzymy ze superrodiny 2 vykazují ATPázovou aktivitu, stimulovanou interakcí s nukleovými kyselinami, ačkoliv ne všechny jsou schopny oddělovat vlákna dvoušroubovice nukleové kyseliny a/nebo po nukleové kyselině translokovat (BYRD a RANEY 2012).

Nejprve byla nalezena sekvenční podobnost mezi K2ORF4p a dvěma helikázami viru *Vaccinia* – nukleosidtrifosfátfosfohydrolázou I (NPH I), kódovanou genem *D11L*, a menší podjednotkou heterodimerního časného transkripčního faktoru viru *Vaccinia* (*Vaccinia virus early transcription factor*, VETF) kódovanou genem *D6R* (WILSON a MEACOCK 1988). O proteinu NPH I je známo, že po stimulaci jednovláknovou DNA poskytuje energii pro elongaci transkripce a pro uvolnění RNA při terminaci transkripce (DENG a SHUMAN 1998). Heterodimerní VETF funguje jako transkripční iniciační faktor, který váže a ohýbá promotorovou oblast DNA časných genů (BROYLES *et al.* 1991). Později byla také nalezena sekvenční podobnost mezi K2ORF4p a homology helikáz D6 a D11 i u dalších nukleocytoplasmatických virů z čeledí *Asfarviridae*, *Phycodnaviridae*, *Iridoviridae* a *Marseilleviridae* (YÁÑEZ *et al.* 1993; LU *et al.* 1996; SATWIKI *et al.* 2012). Sekvenční podobnost K2ORF4p ke zmíněným virovým helikázám je větší než k jiným helikázám ze stejné rodiny, takže se předpokládá společný původ (IYER *et al.* 2001; SATWIKI *et al.* 2012). Přesná role K2ORF4p je zatím neznámá, i když se díky sekvenčním podobnostem očekává role v transkripci lineárních plasmidů. Umístění konzervovaných helikázových motivů K2ORF4p je znázorněno na Obrázku 4.



**Obrázek 4:** Schématické zobrazení sekvence a predikované terciární struktury DExD/H box helicázy lineárních plasmidů kvasinek vykazující podobnost k helicázám ze superrodiny 2. **(A)** Schéma primární struktury K2ORF4p s vyznačenými detekovanými konzervovanými helicázovými motivy (šedé obdélníky) domény 1 (červeně) a domény 2 (modře), popsány podle (GORBALENYA *et al.* 1989; BAYLIS *et al.* 1993; DÜRR *et al.* 2005). **(B)** Model terciární struktury K2ORF4p, pokrývající aminokyselinové zbytky 1–453, zobrazující doménu 1 (červeně) a doménu 2 (modře) helicázového jádra, části neodpovídající helicázovému jádru (světle žlutě), analog ATP (černě) a DNA (oranžově). Model byl vytvořen homologním modelováním serverem Robetta podle Rad54 helicázy (PDB: 1Z3I). N-konec K2ORF4p je označen hvězdičkou. Nukleová kyselina tohoto zobrazení je založena na struktuře Rad54 helicázy v komplexu s DNA (PDB: 1Z63), s níž byl model strukturně porovnán pomocí DaliLite v. 3 a vizualizován pomocí PyMOL 1.3. Zobrazení analogu ATP je založeno na porovnání modelu se strukturou Chd1 helicázy v komplexu s adenosin 5' gama-thiotrifosfátem (PDB: 3MWY).

## 2.2 POXVIRY

Poxviry patří do skupiny nukleocytoplasmatických velkých dvouvláknových DNA virů. Jedná se o skupinu monofyletických eukaryotických virů s velkými genomy, které infikují širokou škálu hostitelů (KOONIN a YUTIN 2010). Protože ne všechny viry této skupiny splňují kritérium vyplývající z názvu, jako například viry čeledi *Poxviridae*, které jsou po celou část životního cyklu v buňce pouze v cytoplasmě (MINNIGAN a MOYER 1985), byl nově navržen název *Megavirales* pro řád, který by všechny tyto monofyletické viry zahrnoval (COLSON *et al.* 2012). Do tohoto řádu patří čeledi *Poxviridae*, *Asfarviridae*, *Phycodnaviridae*, *Iridoviridae*, *Ascoviridae*, *Mimiviridae* a také recentně navržené čeledi jako *Megaviridae*, *Marseilleviridae* nebo *Pandoraviridae* (ARSLAN *et al.* 2011; COLSON *et al.* 2013; PHILIPPE *et al.* 2013). Poxviry se dělí do dvou podčeledí a jsou patogenní pro mnoho druhů savců, ptáků a plazů (podčeď *Chordopoxvirinae*), ale také hmyzu (podčeď *Entomopoxvirinae*). Genom zástupců poxvirů se skládá z jedné lineární molekuly dvouvláknové DNA o velikosti v rozmezí 130–375 tisíců párů bází, kovalentně zakončené vlásenkovou smyčkou na každém konci (MOSS 2013b). Viriony poxvirů jsou obalené lipidovou dvojvrstvou a vykazují komplexní morfologii, kdy maturovaný virion zpravidla tvarem připomíná cihlu nebo vejce (GAYLORD a MELNICK 1953; MORGAN *et al.* 1954). Modelovým systémem pro studium poxvirů se stal virus *Vaccinia*, jenž byl používán jako vakcína proti infekci lidským virem pravých neštovic, ačkoliv přesný původ tohoto viru je nejasný. Může se jednat o vyselektovaný atenuovaný virus pravých neštovic, hybridní virus mezi virem pravých neštovic a kravských neštovic, nebo o virus koňských neštovic, který se již v přírodě nevyskytuje (BAXBY 1999; RAZZELL 1999).

### 2.2.1 Organizace genomu a nomenklatura genů viru *Vaccinia*

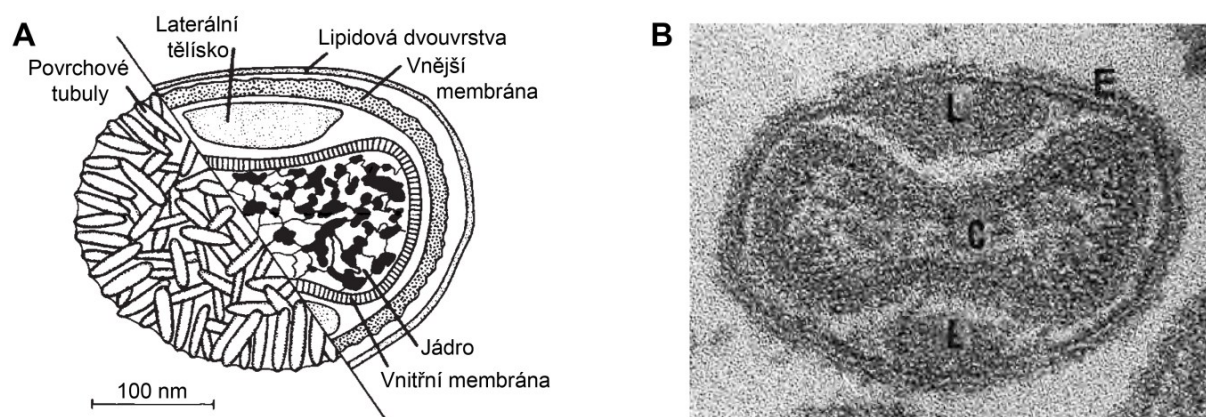
Virus *Vaccinia* (VACV) má genom s vysokým obsahem (~67 %) AT bází o velikosti přibližně 192 tisíc párů bází (GOEBEL *et al.* 1990). Konce genomu obsahují dlouhé invertované terminální repetice (~10 tisíc párů bází), které jsou zakončené invertovanými komplementárními nekompletně spárovanými 101 nt vlásenkami se 4 nt smyčkami, díky nimž je fosfodiesterový řetězec genomu kontinuální viz Obrázek 5 (GARON *et al.* 1978; BAROUDY *et al.* 1982). Invertované terminální repetice obsahují několik genů (WITTEK *et al.* 1980), pro které je tím pádem virus diploidní. Terminální invertované repetice dále obsahují několik krátkých tandemových repetit (WITTEK a MOSS 1980; BAROUDY a MOSS 1982) a vysoce konzervovanou oblast s 20 nt sekvenčním elementem potřebným pro rozdělení konkatemerních forem DNA, vznikajících při replikaci genomu (MERCHLINSKY 1990). Genom kóduje přibližně 200





### 2.2.2 Struktura virionu viru *Vaccinia*

Morfogeneze virionů viru *Vaccinia* je komplexní proces, při kterém vzniká několik forem zralých virionů, které se vzájemně odlišují počtem lipidových dvojvrstev a povrchovými antigeny, pro přehled viz (SMITH *et al.* 2002). Viriony VACV jsou v porovnání s viriony většiny ostatních virů infikujících živočichy poměrně velké a jsou zřetelné i pomocí světelné mikroskopie. Rekonstrukce 3D struktury zralého vnitrobuněčného virionu VACV pomocí kryoelektronové tomografie odhalila cihlovitý tvar virionu o rozměrech 360 x 270 x 250 nm, obaleného jednou lipidovou dvojvrstvou (CYRKLAFF *et al.* 2005). Pod lipidovou dvojvrstvou se nachází vnější membrána, pod kterou jsou zpravidla dvě laterální proteinová tělíska a vnitřní dvojvrstvá membrána bikonkávního jádra virionu ve tvaru činky, které obsahuje genom viru, jenž se koncentruje u vnitřní vrstvy membrány jádra viz Obrázek 6 (CYRKLAFF *et al.* 2005). Střed jádra virionu obsahuje pravděpodobně proteiny potřebné pro syntézu časně mRNA (CYRKLAFF *et al.* 2005). Celková hmotnost virionu VACV je přibližně 10 fg (JOHNSON *et al.* 2006). Díky velikosti a komplexní struktuře obsahují částice VACV velké množství proteinů, které bylo celkem podrobně studováno pomocí hmotnostní spektrometrie. Zralé částice VACV obsahují přibližně 80 proteinů kódovaných virem a 24 proteinů kódovaných hostitelskou buňkou (CHUNG *et al.* 2006; RESCH *et al.* 2007). Jádra virionů se dají samostatně izolovat a obsahují přibližně 50 produktů virových ORF (EASTERBROOK 1966; YODER *et al.* 2006), přičemž 30 produktů jsou enzymy a polovina z nich se přímo podílí na biogenezi časně virové mRNA [shrnuť v (MOSS 2013b)]. Jádra virionů jsou schopná produkce mRNA *in vitro* (KATES a BEESON 1970a). Laterální tělíska se po vstupu viru do buňky rozvolňují a uvolňují například protein bránící protivirové obraně přes interferon  $\gamma$  (SCHMIDT *et al.* 2013).



**Obrázek 6:** (A) Schématické znázornění typického cihlovitého zralého virionu viru *Vaccinia*. Genomová DNA viru a proteiny v jádře jsou organizovány jako nukleosom. Jádro má membránu o tloušťce 9 nm. Laterální tělíska jsou připojena na vnější membránu o tloušťce 12 nm. Povrch vnější membrány se sestává z nepravidelně tvarovaných tubulů. (B) Elektronmikroskopický obraz tenkého řezu zralým virionem viru *Vaccinia*, ukazující jádro virionu (C), laterální tělíska (L) a vnější membránu (E). Upraveno podle (DAMON 2013; MOSS 2013b).

### 2.2.3 Životní cyklus viru *Vaccinia*

Stejně jako mnoho ostatních virů poxviry koordinují replikaci genomu a skládání virionů pomocí časové regulace exprese jednotlivých genů. Životní cyklus VACV začíná vstupem viru do buňky přes plasmatickou membránu pomocí fúze membrán nebo endocytózy. VACV kóduje 15 proteinů, zodpovědných za uchycení viru na plasmatickou membránu a vstup viru do buňky [shrnutí v (MOSS 2016)]. Následuje uvolnění virové částice do cytoplasmy, kde se částečně rozvolní její jádro, které začne po 20 minutách pomocí enkapsidovaného transkripčního aparátu produkovat mRNA přibližně 110 časných genů, která je následně translatována na polysomech (KATES a MCAUSLAN 1967; HOLOWCZAK 1972; METZ *et al.* 1975; BALDICK a MOSS 1993; YANG *et al.* 2010). Mezi produkty časných genů patří proteiny podílející se na replikaci DNA, transkripci středních genů a úniku imunitní obraně hostitele (JONES a MOSS 1984; MOORE a SMITH 1992; SANZ a MOSS 1999). Jádra virových částic VACV jsou transportována pomocí mikrotubulů do specifických cytoplasmatických míst, kde dochází k dalšímu rozvolnění jádra částice a tvorbě virových továren, ve kterých probíhá replikace DNA uvolněné z jádra částice a transkripce a translace 53 středních a 38 pozdních genů (HARFORD *et al.* 1966; TOLONEN *et al.* 2001; CARTER *et al.* 2003; KATSAFANAS a MOSS 2007; YANG *et al.* 2011b). Mezi produkty středních genů, jejichž mRNA se objevuje ~100 minut po infekci, patří především proteiny podílející se na vazbě a pakázování DNA a transkripci pozdních genů (KECK *et al.* 1990; BALDICK a MOSS 1993; YANG *et al.* 2011b). Produkty pozdních genů, jejichž mRNA se objevuje ~140 minut po infekci, jsou proteiny pro morfogenezi částic, membránové proteiny zralých částic a transkripční faktory časných genů (GERSHON a MOSS 1990; BALDICK a MOSS 1993; YANG *et al.* 2011b).

Sestavování virionů začíná tvorbou půlměsíčitých membránových struktur, patrně odvozených od endoplasmatického retikula, skládajících se v nezralé částice, do kterých je pakázována genomová DNA [shrnutí v (MOSS 2015)]. Nezralé částice jsou sférické a jejich vnější mřížovaný povrch je tvořen trimery *scaffold* proteinu D13 (SZAJNER *et al.* 2005). Nezralé částice maturují ve zralé částice metamorfózou, která vyžaduje proteolytické štěpení několika hlavních strukturních proteinů virionu (BYRD *et al.* 2002). Některé zralé částice jsou obaleny dalšími dvěma lipidovými dvojvrstvami, obsahujícími 9 virových proteinů [shrnutí v (MOSS 2013b)]. Toto obalení je závislé na retrográdním transportu z časných endozómů do *trans* Golgiho aparátu (SIVAN *et al.* 2016). Obalené částice jsou transportovány k periférii buňky a exocytovány. Transport probíhá po mikrotubulech (HOLLINSHEAD *et al.* 2001), navzdory přechozímu zjištění, že je k tomuto účelu využíván aktin (CUDMORE *et al.* 1995). Část obalených virionů zůstane uchycených k buněčné membráně a indukuje nukleaci aktinu pod membránou, což usnadňuje rozšíření viru do sousedících buněk (CUDMORE *et al.* 1995; SMITH a LAW 2004).

Většina zralých virionů v buňce však zůstane neobalena dalšími dvěma lipidovými dvojvrstvami a může být uvolněna buněčnou lyzí infikované buňky.

#### 2.2.4 Replikace DNA viru *Vaccinia*

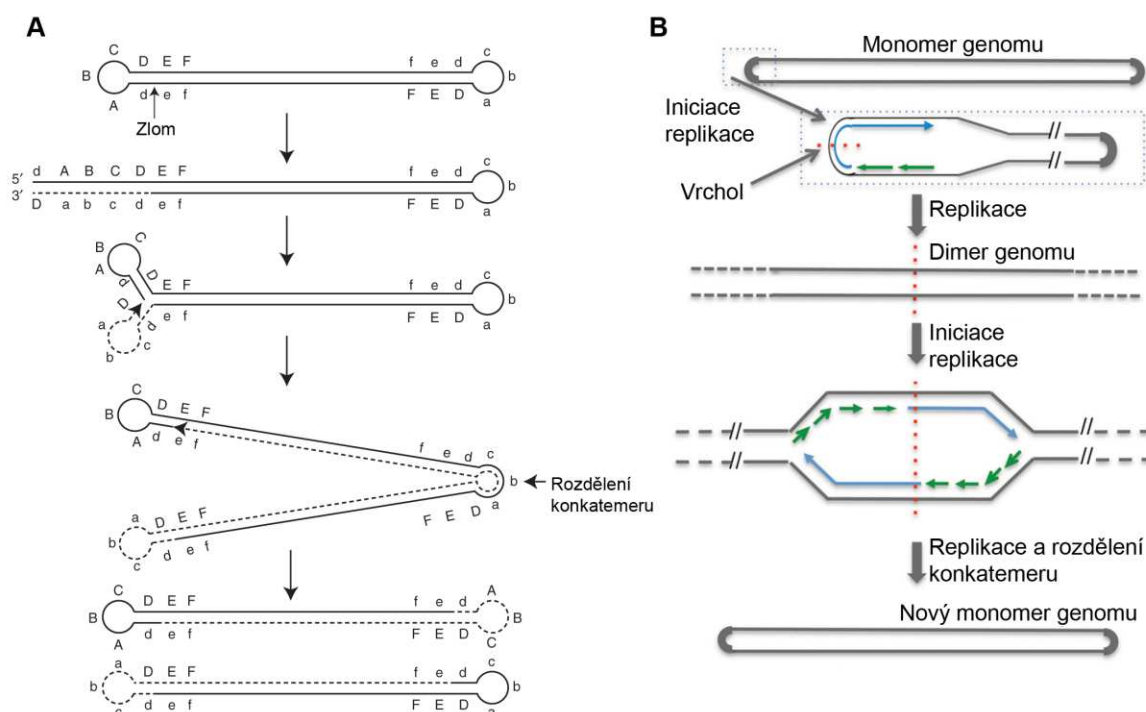
Na rozdíl od většiny DNA virů se poxviry replikují v cytoplasmě a ne v jádru hostitelské buňky a dokáží se replikovat i v bezjaderných buňkách (PRESCOTT *et al.* 1971). Díky tomu vykazují poxviry značnou autonomii. Původní předpoklad, že jejich replikace a transkripce je úplně nezávislá na buněčných proteinových faktorech, se nepotvrdil. V dnešní době je však známo několik buněčných proteinů, které VACV při své replikaci a transkripci využívá, jako třeba DNA ligáza I a TATA-box vazebný protein (TBP) (KNUTSON *et al.* 2006; PARAN *et al.* 2009). Všechny nebo téměř všechny proteiny potřebné k replikaci, úpravě a pakázování DNA viru *Vaccinia*, jsou kódované virem, přičemž pět proteinů je pro samotnou replikaci DNA VACV esenciálních [shrnutí v (MOSS 2013a)]. Jedná se o DNA polymerázu, procesivní faktor DNA polymerázy, fúzní primázu/helikázu, *single stranded DNA binding* protein a uracil DNA glykosylázu (SRIDHAR a CONDIT 1983; MILLNS *et al.* 1994; EVANS *et al.* 1995; ROCHESTER a TRAKTMAN 1998; PUNJABI *et al.* 2001; DE SILVA *et al.* 2007).

K začátku replikace DNA VACV dochází zhruba 1–2 hodiny po infekci ve virových továrnách, obklopených membránou endoplasmatického retikula, přičemž z jednoho virionu může vzniknout jedna virová továrna s tím, že se jednotlivé virové továrny postupně spojují dohromady (CAIRNS 1960; JUNGWIRTH a LAUNER 1968; TOLONEN *et al.* 2001; KATSAFANAS a MOSS 2007; LIN a EVANS 2010). Díky kovalentně uzavřeným koncům genomu poxvirů vznikají při replikaci DNA přechodné dimerní a tetramerní konkatemery, v nichž jsou jednotlivé genomy orientovány zrcadlově (MOYER a GRAVES 1981). Konkatemery jsou před inkorporací do virových částic VACV štěpeny na monomery genomu pomocí esenciální resolvázy Hollidayovy struktury, kódované virem (GARCIA a MOSS 2001). Původní experimenty naznačily, že k počátku replikace DNA dochází poblíž konců genomu VACV, nicméně existence definovaného počátku replikace DNA byla zpochybněna, protože jakýkoliv cirkulární DNA plasmid transfekovaný do buněk infikovaných poxviru se replikuje v cytoplasmatických virových továrnách způsobem, závislým na virových proteinech esenciálních pro replikaci DNA (POGO *et al.* 1984; DELANGE a MCFADDEN 1986; DE SILVA a MOSS 2005). Optimální efektivita replikace lineárních DNA minichromosomů s terminálními vlásenkami VACV sice vyžaduje přítomnost přibližně 200 párů bází dlouhé sekvence odvozené od konce genomu VACV, tato oblast však obsahuje známý element potřebný pro štěpení konkaterů (DU a TRAKTMAN 1996). Nová studie se zaměřila na hledání počátku replikace VACV pomocí hlubokého sekvenování a jako



nejvíc zastoupený iniciační bod byl identifikován G nukleotid v terminální vlásence na jednom z konců genomu viz Obrázek 5, a stejný nukleotid na rozhraní konkatemerního spojení replikačních intermediátů (SENKEVICH *et al.* 2015).

Bylo představeno několik modelů replikace DNA VACV. Díky přítomnosti DNA vlásenek na konci genomu VACV byl až donedávna upřednostňován model otáčející se vlásenky s vytěsňováním vlákna, jako u virů z čeledi *Parvoviridae* (BAROUDY *et al.* 1982; MOSS 2013a). V tomto modelu dochází k jednovláknovému štěpení DNA poblíž terminální vlásenky, vzniku volné 3' OH skupiny a kontinuálnímu prodlužování řetězce, protože díky komplementaritě bází v rámci vlásenky dojde k překlopení syntetizovaného vlákna viz Obrázek 7A (MOYER a GRAVES 1981; BAROUDY *et al.* 1982). V dnešní době se však pohled na replikaci DNA VACV mění. Je to díky tomu, že byly u VACV identifikovány předpokládané Okazakiho fragmenty, které nasvědčují diskontinuální syntéze opoždujícího se vlákna (POGO a O'SHEA 1978). Všechny známé poxviry také kódují fúzní primázu/helikázu, která je pro replikaci DNA esenciální, což naznačuje používání RNA primerů při replikaci DNA (DE SILVA *et al.* 2007). V novém modelu replikace DNA tak dochází k iniciaci replikace v rámci vlásenky a semidiskontinuální syntéze s využitím vedoucího i opoždujícího se vlákna viz Obrázek 7B (SENKEVICH *et al.* 2015). Na rozhraní zrcadlově orientovaného dimeru pak může docházet k další iniciaci replikace DNA, která je dvousměrná (SENKEVICH *et al.* 2015).



**Obrázek 7:** Modely replikace DNA u poxvirů. (A) Model jednovláknového zlomu a vytěsňování vlákna. Přerušované čáry znázorňují nově syntetizované vlákno DNA a šipky značí 3' OH konec. Komplementární sekvence jsou zobrazeny malými a velkými písmeny. (B) Model využívající RNA primery a syntézu vedoucího (modré šipky) a opoždujícího se (zelené šipky) vlákna. Vrchol spojení konkateru zobrazen červenými pomlčkami. Upraveno podle (MOSS 2013a; SENKEVICH *et al.* 2015).

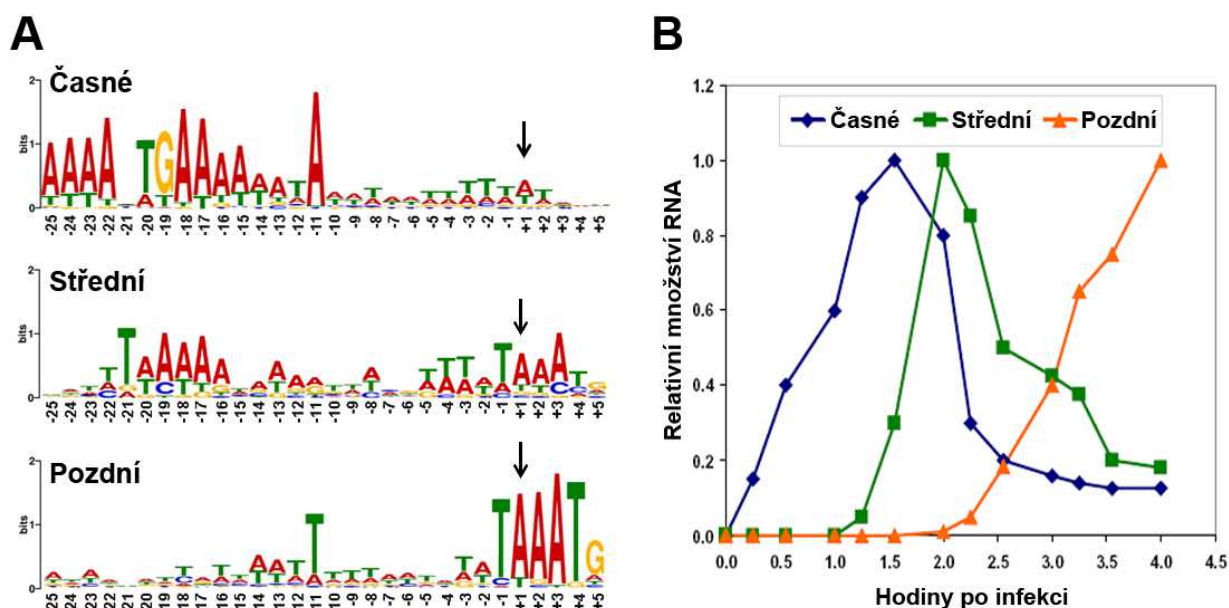
### 2.2.5 Regulace genové exprese a transkripce viru *Vaccinia*

Genová exprese viru *Vaccinia* během infekce probíhá časovou kaskádou zahrnující sekvenční syntézu transkriptů časných, středních a pozdních genů viz Obrázek 8B a je regulována především na úrovni iniciace transkripce (BALDICK a MOSS 1993; BROYLES 2003). Každá ze tří časových tříd genů má specifické promotory a vyžaduje rozdílné *trans* regulující transkripční faktory, kdy každá genová třída kóduje regulační faktory potřebné pro transkripci následující genové třídy, čímž je zajištěna sekvenční posloupnost exprese (GERSHON a MOSS 1990; KECK *et al.* 1990; KOVACS *et al.* 1994; SANZ a MOSS 1999).

Jádro částice VACV obsahuje transkripční systém schopný produkce mRNA časných genů, která je vysunována ven přes předpokládané póry vnitřní membrány částice po částečném rozvolnění částice při infekci (KATES a MCAUSLAN 1967; HOLOWCZAK 1972; METZ *et al.* 1975; GOLINI a KATES 1985; ROHRMANN a MOSS 1985). Vysunování časně mRNA z jádra částice je závislé na ATP (KATES a BEESON 1970a). Časná mRNA v buňce asociuje s mikrotubuly a je akumulována v amorfních cytoplasmatických granulárních strukturách, obklopených polysomy (METZ *et al.* 1975; MALLARDO *et al.* 2001). Transkripce středních a pozdních genů VACV vyžaduje syntézu nové RNA polymerázy a je vázána na replikaci DNA, neboť inhibitory replikace DNA inhibují transkripci středních a pozdních genů (ODA a JOKLIK 1967; VOS a STUNNENBERG 1988; HOODA-DHINGRA *et al.* 1989). K transkripci středních a pozdních genů dochází ve virových továrnách a je k ní vyžadována přítomnost nahého DNA templátu (KECK *et al.* 1990; KATSAFANAS a MOSS 2007). Střední a pozdní geny se někdy souhrnně označují jako postreplikativní geny a vzájemné rozlišení genů těchto dvou tříd bylo poměrně obtížné, až nedávno došlo k sestavení první časové transkripční mapy VACV (YANG *et al.* 2011b).

Všechny geny VACV zřejmě obsahují vlastní promotory, které jsou, nezávisle na časové třídě, poměrně krátké a jednoduché, neboť jsou lokalizovány v přibližně 30 nt dlouhé sekvenci proti směru transkripce od transkripčního iniciačního místa viz Obrázek 8A (DAVISON a MOSS 1989a, b). U některých genů byla popsána kontinuální exprese během celého infekčního cyklu, což je způsobeno tandemovým uspořádáním promotorů různých časových tříd nebo kompozitním charakterem daného promotoru (WITTEK *et al.* 1980; COCHRAN *et al.* 1985; BROYLES a MOSS 1986; AHN *et al.* 1990a). Bylo také ukázáno, že jeden časný gen je přepisován v pozdní době infekčního cyklu, a že mnoho středních promotorů vykazuje i pozdní aktivitu (GARCÉS *et al.* 1993; YANG *et al.* 2013). Transkripční faktory a promotory jsou konzervovány, protože promotory jednoho druhu poxviru jsou funkční v buňce infikované jiným poxvirem (KUMAR a BOYLE 1990). V následujících kapitolách budou nejdříve popsány enzymy a reakce

společné pro syntézu a modifikace mRNA všech časových tříd genů a potom budou uvedeny rozdíly v transkripci mezi jednotlivými třídami genů.



**Obrázek 8:** (A) Sekvenční motivy promotorů časných, středních a pozdních genů viru *Vaccinia*. Motivů byly vygenerovány programem MEME na základě celogenomových analýz transkriptů pomocí hlubokého sekvenování (YANG *et al.* 2011a; YANG *et al.* 2011b). Šipka udává místo iniciace transkripce. (B) Hladiny transkriptů časných (modře), středních (zeleně) a pozdních (oranžově) genů viru *Vaccinia* v HeLa buňkách. K detekci a měření mRNA v ustáleném stavu byly použity RNA sondy, specifické pro 5' konce mRNA kódující *C11R* (časný), *G8R* (střední) a *F17R* (pozdní) otevřený čtecí rámec, hybridizace a ochrana před štěpením RNázou (BALDICK a MOSS 1993). Maximální hladina RNA pro každý gen je nastavena na hodnotu 1.0. Upraveno podle (MOSS 2013b).

### 2.2.5.1 RNA polymeráza

RNA polymeráza kódovaná VACV je vícepodjednotkový enzym, který existuje ve dvou formách, přičemž jedna forma je specifická pro transkripci časných genů a druhá forma pro transkripci postreplicative genů (BAROUDY a MOSS 1980; AHN a MOSS 1992; WRIGHT a CORONEOS 1995). Obě formy RNAP mají 8 společných podjednotek a odlišují se pouze tím, že RNAP přepisující časné geny obsahuje protein RAP94 (*RNA polymerase-associated protein 94 kDa*), který RNAP udílí specifitu k promotorům a terminátorům časných genů (AHN *et al.* 1994; YANG a MOSS 2009). Z osmi společných podjednotek RNAP sdílejí dvě největší podjednotky Rpo147 a Rpo132 rozsáhlou sekvenční podobnost s největšími podjednotkami ( $\beta'$  a  $\beta$  u bakterií) vícepodjednotkových RNA polymeráz (BROYLES a MOSS 1986; PATEL a PICKUP 1989). U dalších čtyř podjednotek RNAP VACV byla nalezena sekvenční podobnost k přídatným podjednotkám eukaryotických RNAP viz Tabulka 1, a podjednotka Rpo30 má sekvenční podobnost s elongačním faktorem TFIIS eukaryotické RNA polymerázy II (AHN *et al.*

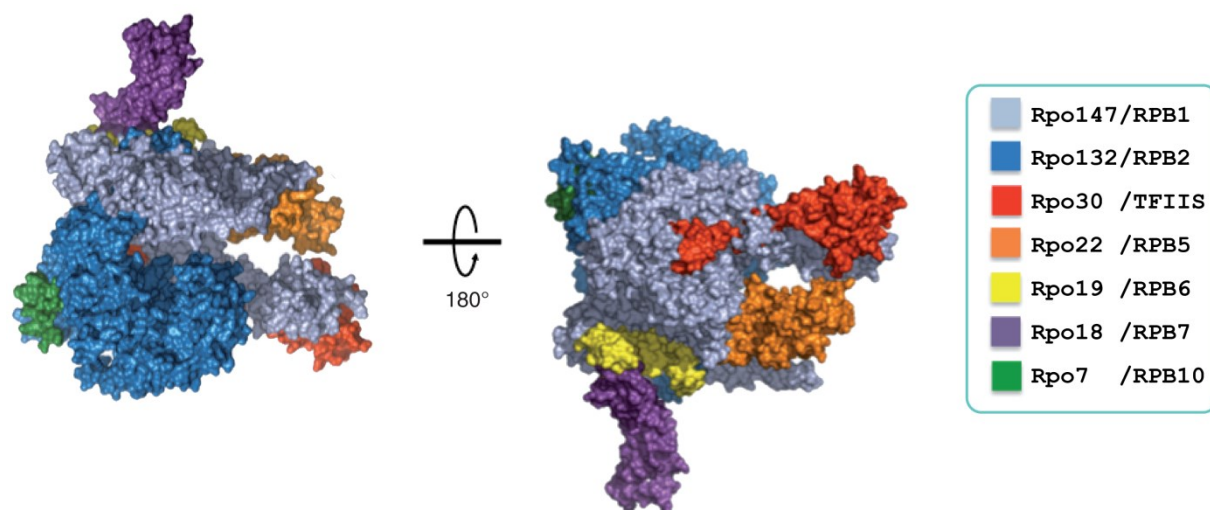
1990a; AMEGADZIE *et al.* 1992a; KNUTSON a BROYLES 2008). Pouze jediná podjednotka RNAP VACV Rpo35 nevykazuje žádnou signifikantní podobnost s proteiny, jejichž sekvence jsou dostupné v databázích (KNUTSON a BROYLES 2008).

Tabulka 1: RNA polymeráza viru <i>Vaccinia</i>			
Protein	Gen	Buněčné homology / Funkce	Reference
Rpo147	<i>J6R</i>	RPB1 ( $\beta'$ u bakterií)	BROYLES a MOSS (1986)
Rpo132	<i>A24R</i>	RPB2 ( $\beta$ u bakterií)	PATEL a PICKUP (1989)
Rpo35	<i>A29R</i>		AMEGADZIE <i>et al.</i> (1991)
Rpo30	<i>E4L</i>	TFIIS	AHN <i>et al.</i> (1990a)
Rpo22	<i>J4R</i>	RPB5	BROYLES a MOSS (1986)
Rpo19	<i>A5R</i>	RPB6 ( $\omega$ u bakterií)	AHN <i>et al.</i> (1992)
Rpo18	<i>D7R</i>	RPB7	AHN <i>et al.</i> (1990b)
Rpo7	<i>G5.5R</i>	RPB10	AMEGADZIE <i>et al.</i> (1992a)
RAP94	<i>H4L</i>	Specifita časné transkripce	AHN a MOSS (1992)

Celková velikost komplexu RNA polymerázy VACV je přibližně 500 kDa a díky sekvenčním podobnostem většiny podjednotek RNAP VACV s podjednotkami eukaryotických RNAP lze vytvořit téměř kompletní model komplexu RNA polymerázy VACV (Obrázek 9). Podjednotky Rpo147 a Rpo132 obsahují všechny konzervované oblasti katalytických podjednotek buněčných RNAP ( $\beta'$  a  $\beta$  u bakterií), a proto jsou RNAP VACV řazeny ke kanonickým vícepodjednotkovým RNA polymerázám (LANE a DARST 2010a). RNA polymeráza VACV však chybí zjevné orthology  $\alpha$  podjednotek, což by značilo výraznou změnu v architektuře a uspořádání jádra enzymu. Zároveň se však nedá vyloučit, že roli těchto podjednotek zajišťuje podjednotka Rpo35 (KNUTSON a BROYLES 2008). Virová RNAP je v purifikovaném stavu rezistentní k  $\alpha$ -amanitinu a rifampicinu a schopná syntézy RNA na nespecifických jednovláknových DNA templátech za přítomnosti manganatých iontů (BAROUDY a MOSS 1980; SPENCER *et al.* 1980). Purifikovaná RNAP VACV také vykazuje opravnou štěpící aktivitu, ke které bude nejspíš přispívat podjednotka Rpo30 (HAGLER a SHUMAN 1993). Analýza elongačních komplexů RNA polymerázy VACV odhalila obdobnou délku transkripční bubliny a DNA:RNA hybridní molekuly jako u elongačních komplexů ostatních vícepodjednotkových RNA polymeráz, takže se nejspíš jedná o konzervovaný funkční prvek těchto enzymů (HAGLER a SHUMAN 1992b). Elongační komplexy RNAP VACV jsou také poměrně stabilní vůči působení soli, detergentů a kompetitorové DNA (HAGLER a SHUMAN 1992c).

Geny pro RNAP nukleocytoplasmatických virů pochází patrně z genů kódujících RNAP ancestrálních eukaryotických hostitelů těchto virů (SONNTAG a DARAI 1995; IYER *et al.* 2006). Na základě srovnání aminokyselinových sekvencí konzervovaných domén orthologů  $\beta$  a  $\beta'$

podjednotek vícepodjednotkových RNAP bylo pozorováno, že RNAP virů čeledi *Poxviridae* se jeví fylogeneticky nejpříbuznější eukaryotické RNA polymeráze I (LANE a DARST 2010a). Zdá se, že ortholog  $\beta'$  podjednotky původní RNAP společného předka všech nukleocytoplasmatických virů má původ právě v eukaryotické RNA polymeráze I (YUTIN a KOONIN 2012).



**Obrázek 9:** Horní a spodní pohled na prostorové zobrazení predikované terciární struktury RNA polymerázy viru *Vaccinia*, vykazující sekvenční podobnost k eukaryotickým RNA polymerázám. Jednotlivé podjednotky RNAP jsou barevně odlišeny a popsány v legendě na pravé straně obrázku, včetně názvů podjednotek eukaryotické RNA polymerázy II a TFIIIS, podle kterých byly homologně modelovány. Upraveno podle (BROYLES a KNUTSON 2010).

### 2.2.5.2 Posttranskripční modifikace mRNA

Transkripty genů všech časových tříd VACV podléhají úpravám díky RNA modifikačním enzymům kódovaným virem (Tabulka 2). Transkripty časných i postreplikativních genů obsahují 5' čepičkovou strukturu typu *cap* 1 (URUSHIBARA *et al.* 1975; WEI a MOSS 1975; BOONE a MOSS 1977). Čepičková struktura typu *cap* 1 obsahuje kromě N<sup>7</sup>-methylguanositinu připojeného 5'-5' trifosfátovou vazbou na 5' konec mRNA (typ *cap* 0) ještě methylovou skupinu na 2' hydroxylové skupině ribózového zbytku prvního nukleotidu 5' konce mRNA. Za syntézu čepičkové struktury transkriptů VACV jsou zodpovědné dva virové enzymy - heterodimerní mRNA *capping* enzym složený z podjednotek D1 a D12 syntetizuje na mRNA strukturu typu *cap* 0, a mRNA *cap* (nukleosid-2'-O)-methyltransferáza J3 následně modifikuje mRNA nesoucí *cap* 0 na *cap* 1 (BARBOSA a MOSS 1978; VENKATESAN *et al.* 1980; MORGAN *et al.* 1984; NILES *et al.* 1989; SCHNIERLE *et al.* 1992). Transkripty časných i postreplikativních genů VACV, izolované z HeLa buněk, obsahují *in vivo* ještě další methylové skupiny, jmenovitě na bázi prvního nukleotidu mRNA (modifikovaný typ *cap* 1) a 2' hydroxylové skupině ribózového zbytku druhého nukleotidu mRNA (*cap* 2), které jsou typické pro buněčné mRNA vyšších eukaryot (SHATKIN 1976; BOONE a MOSS 1977). Na těchto dalších methylacích mRNA VACV se zřejmě podílí



buněčné enzymy, neboť buněčná mRNA *cap* (nukleosid-2'-O)-methyltransferáza typu *cap* 2 se nachází v cytoplasmatické frakci lyzátu neinfikovaných HeLa buněk (LANGBERG a MOSS 1981).

Neexistují důkazy, že by 5' konce transkriptů VACV vznikaly exo- nebo endonukleolytickým štěpením, jelikož k připojení čepičkové struktury u časných mRNA dochází na iniciačním nukleotidu transkripce (VENKATESAN a MOSS 1981). U genů VACV také nedochází k sestřihu RNA, protože u časných i postreplikativních genů nebyly detekovány intervenující sekvence štěpením DNA:RNA hybridních molekul S1 endonukleázou a genom kóduje otevřené čtecí rámce (WITTEK *et al.* 1980; COOPER *et al.* 1981; GOEBEL *et al.* 1990). Tvorba 3' konců transkriptů časných genů VACV probíhá specifickým mechanismem terminace transkripce a neexistují důkazy, že by 3' konce vznikaly exo- nebo endonukleolytickým štěpením (ROHRMANN *et al.* 1986; SHUMAN *et al.* 1987). Tvorba 3' konců transkriptů postreplikativních genů VACV není dopodrobna prozkoumána hlavně díky velké heterogenitě délky 3' konců jednotlivých transkriptů daného genu, což znemožňuje mapování konců pomocí S1 endonukleázy (COOPER *et al.* 1981; MAHR a ROBERTS 1984a; WEIR a MOSS 1984) a vede k detekci široce rozprostřeného signálu při detekci postreplikativních transkriptů metodou *northern blot* (AHN *et al.* 1992; XIANG *et al.* 1998). Zřejmě díky rozdílné délce 3' konců se transkripty jednotlivých postreplikativních genů VACV často překrývají (COLBY 1971; BOONE a MOSS 1978). Definované 3' konce byly u postreplikativních transkriptů zatím detekovány jen u genů *A24R* (kontinuální exprese), *A25L* (střední) a *F17R* (pozdní) VACV, homologních genů u jiných poxvirů, a u postreplikativních transkriptů prepisovaných z promotoru poblíž konce genomu VACV (PATEL a PICKUP 1989; PARSONS a PICKUP 1990; AMEGADZIE *et al.* 1992b; ANTCHAK *et al.* 1992; D'COSTA *et al.* 2004). U transkriptů genů *A25L* a *F17R* jsou 3' konce generovány endonukleolytickým štěpením a nukleázová aktivita je spojena s proteinovým komplexem obsahujícím virový protein H5 (ANTCHAK *et al.* 1992; D'COSTA *et al.* 2004; D'COSTA *et al.* 2008). U transkriptu *A25L* byla identifikována specifická dvousložková *cis* regulační 43 nt dlouhá sekvence, která je pro štěpení 3' konce RNA nutná (HOWARD *et al.* 1999). Zatím však není jasné, jestli je úprava 3' konců postreplikativních transkriptů u VACV běžná nebo výjimečná (BROYLES 2003).

**Tabulka 2:** Faktory posttranskripčních modifikací mRNA viru *Vaccinia*

Faktor	Gen	kDa	Funkce	Reference
Capping enzym	<i>D1R</i> , <i>D12L</i>	97, 33	5' mRNA capping ( <i>cap</i> 0)	MORGAN <i>et al.</i> (1984) NILES <i>et al.</i> (1989)
(nukleosid-2'-O)-methyltransferáza	<i>J3R</i>	39	5' mRNA capping ( <i>cap</i> 1)	SCHNIERLE <i>et al.</i> (1992)
Štěpící faktor	<i>H5R</i>	22	3' mRNA štěpení	D'COSTA <i>et al.</i> (2008)
Poly(A) polymeráza	<i>E1L</i> , <i>J3R</i>	55, 39	3' mRNA polyadenylace	GERSHON <i>et al.</i> (1991)

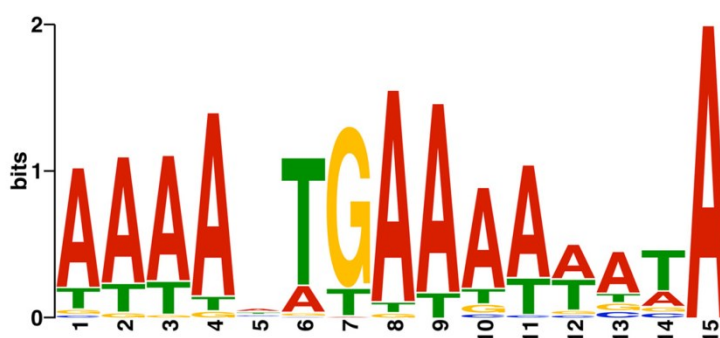
Transkripty genů všech časových tříd VACV obsahují 3' poly(A) konec o průměrné délce přibližně 100 nt *in vivo* (KATES a BEESON 1970b; NEVINS a JOKLIK 1975). Za syntézu 3' poly(A) konce transkriptů VACV je zodpovědná virová poly(A) polymeráza (PAP) složená z podjednotek E1 a J3 (MOSS *et al.* 1975; GERSHON *et al.* 1991). Větší podjednotka PAP E1 (známá též jako VP55) je katalytická a dokáže procesivně přidávat pouze 30–35 adenylátů na 3' konec RNA *in vitro*, menší podjednotka PAP J3 (VP39) váže 3' poly(A) konec RNA a stimuluje větší podjednotku PAP k semiprocessivnímu přidávání dalších adenylátových zbytků do celkové délky 200–300 nukleotidů *in vitro* (GERSHON a MOSS 1992, 1993a). Stimulativní podjednotka PAP je zároveň mRNA *cap* (nukleosid-2'-O)-methyltransferázou, ačkoliv tyto dvě aktivity jsou na sobě nezávislé (GERSHON *et al.* 1998). Pro vazbu E1 podjednotky PAP k RNA templátu *in vitro* je potřeba 5'-U<sub>2</sub>N<sub>15</sub>U-3' motiv vzdálený 10 nt od 3' konce RNA (GERSHON a MOSS 1993b; DENG *et al.* 1997; DENG a GERSHON 1997). Pro syntézu 3' poly(A) konce RNA delšího než 3–7 nt *in vitro* vyžaduje E1 podjednotka kromě tří uridinů výše zmíněného vazebného motivu ještě přítomnost šesti uridinů 5'-U<sub>5</sub>N<sub>1</sub>U-3' motivu na 3' konci templátové RNA (JOHNSON *et al.* 2004). Kromě bezprostředního okolí katalytického centra nevykazuje PAP VACV strukturní podobnost s jednopodjednotkovými eukaryotickými poly(A) polymerázami (MOURE *et al.* 2006). Specifické sekvence na 3' koncích mRNA však možná nemusí být pro 3' polyadenylaci nutné, protože předčasně terminované transkripty VACV *in vitro* a *in vivo* obsahují 3' poly(A) konec (GERSHOWITZ a MOSS 1979; XIANG *et al.* 2000).

## 2.2.6 Transkripce časných genů viru *Vaccinia*

Všechny enzymy potřebné pro syntézu a úpravu mRNA časných genů VACV se nacházejí v jádře částice, které cca 20 minut po infekci začne produkovat mRNA přibližně poloviny genů s nejvyšší detekovanou hladinou časné mRNA kolem 100 minut po synchronní infekci HeLa buněk (BALDICK a MOSS 1993; YANG *et al.* 2010). Kromě virové RNAP asociované s proteinem RAP94 jsou potřeba jen tři další proteinové faktory pro iniciaci a terminaci transkripce časných genů *in vitro* (Tabulka 3). Virová RNAP postrádající protein RAP94 nedokáže iniciovat transkripci na templatech s časným promotorem *in vitro* (AHN *et al.* 1994).

Tabulka 3: Časné transkripční faktory viru <i>Vaccinia</i>				
Faktor	Gen	kDa	Funkce	Reference
VETF	A7L, D6R	82, 70	Iniciace transkripce	GERSHON a MOSS (1990) BROYLES a FESLER (1990)
VTF	D1R, D12L	97, 33	Terminace transkripce/ Capping enzym	SHUMAN <i>et al.</i> (1987)
NPH I	D11L	72	Terminace transkripce	DENG a SHUMAN (1998)

Promotory časných genů VACV jsou relativně krátké a umístěné proti směru transkripce (VASSEF *et al.* 1985; MARS a BEAUD 1987). Rozsáhlé mutační analýzy vybraného časného promotoru identifikovaly esenciální 16 nukleotidové AT bohaté jádro promotoru nazvané Kontrolní element proti směru transkripce (*Upstream control element*; UCE), které je umístěno přibližně 14 nt proti směru transkripce od transkripčního iniciačního místa a obsahuje konzervovaný G nukleotid (DAVISON a MOSS 1989a). Novými metodami hlubokého sekvenování transkriptomu VACV byla určena 15 nukleotidová konsenzus sekvence (Obrázek 10), shodná s dříve detekovanou sekvencí UCE a lokalizovaná v průměrné vzdálenosti 12 nt před iniciačním nukleotidem transkripce časných genů (YANG *et al.* 2011a). Nukleotidová sekvence UCE a míra její podobnosti s konsenzuální sekvencí ovlivňuje hladinu transkripce reportérového genu (DAVISON a MOSS 1989a). K iniciaci transkripce časných genů *in vivo* je však kromě UCE zřejmě potřeba ještě vysoký obsah AT bází v oblasti mezi UCE a iniciačním nukleotidem transkripce (YANG *et al.* 2011a). Iniciačním nukleotidem transkripce *in vivo* i *in vitro* bývá zpravidla purin (BOONE a MOSS 1977; KEITH *et al.* 1980; YANG *et al.* 2011a). V transkripčním systému *in vitro* vede mutace purinového transkripčního iniciačního nukleotidu za pyrimidinový nukleotid k posunu iniciačního místa transkripce po nebo proti směru transkripce na nejbližší purinový nukleotid (DAVISON a MOSS 1989a). Časné mRNA mají obecně poměrně krátké 5' nepřekládané oblasti, jak bylo detekováno metodou prodlužování primeru a S1 endonukleázovým mapováním DNA:RNA hybridních molekul (WEIR a MOSS 1983; MAHR a ROBERTS 1984b; LEE-CHEN *et al.* 1988). Průměrná délka 5' nepřekládaných oblastí časných mRNA je 22 nukleotidů včetně iniciačního nukleotidu (YANG *et al.* 2011a).



**Obrázek 10:** Konsenzus sekvence UCE motivu promotorů časných genů viru *Vaccinia* odvozená pomocí programu MEME ze sekvencí o délce 50 nt, předcházejících transkripční iniciační místa 84 časných otevřených čtecích rámců, jež byla identifikována na základě celogenomových analýz transkriptů VACV pomocí hlubokého sekvenování. Konzervovaný guanosinový nukleotid se nachází na pozici 7. Upraveno podle (YANG *et al.* 2011a).

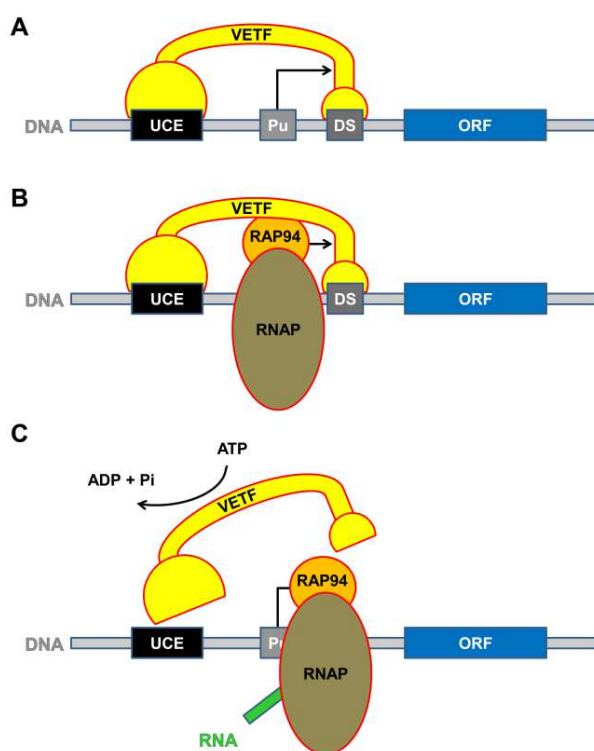


Před 3' konci časných mRNA byl identifikován výskyt opakujících se krátkých úseků T nukleotidů a následně pomocí transkripčního systému *in vitro* byl nalezen krátký heptanukleotidový signál 5'-TTT<sup>\*</sup>TTNT-3', vyskytující se zhruba 50 nt proti směru transkripce od 3' konce mRNA, který je nutný a dostačující pro terminaci transkripce (YUEN a MOSS 1986, 1987). Tvorba 3' konců transkriptů časných genů VACV probíhá specifickým mechanismem terminace transkripce, protože transkripty plné délky vznikají poměrně rychle v *in vitro* transkripčním systému a transkripty delší detekovány nebyly (ROHRMANN *et al.* 1986). Delší transkripty připravené *in vitro* pomocí T7 RNA polymerázy po inkubaci s extrakty VACV nejsou upravovány, což hovoří proti vzniku 3' konců mRNA exo- nebo endonukleolytickým štěpením (ROHRMANN *et al.* 1986). Signál 5'-TTT<sup>\*</sup>TTNT-3' se zpravidla nachází pouze poblíž 3' konců časných ORF a ne v kódující oblasti genu, i když existují i výjimky (YUEN a MOSS 1987; LEE-CHEN *et al.* 1988; PATEL a PICKUP 1989). Výskyt terminačního signálu v cizím genu, exprimovaném pod kontrolou časného promotoru z genomu VACV, vede k předčasné terminaci transkripce tohoto genu (EARL *et al.* 1990). Bylo naznačeno, že terminační signál transkripce není 100% účinný, neboť *in vivo* vznikají i delší časné mRNA, a efektivita samostatně se vyskytujícího signálu byla určena na 75–80 % (BAJSZÁR *et al.* 1983; EARL *et al.* 1990). Objevující se překryvné nebo tandemově uspořádané terminační signály genů tak mohou zvyšovat celkovou efektivitu terminace transkripce *in vivo*. Tento specifický mechanismus terminace transkripce je omezen pouze na molekuly RNAP, které rozpoznávají časné promotory (CONDIT *et al.* 1996a).

Další analýzy využívající inkorporaci bromo- a iodouridinu do RNA během transkripce *in vitro* ukázaly, že terminační signál je rozpoznáván v podobě 5'-UUUUUNU-3' v nascentní mRNA (SHUMAN a MOSS 1988, 1989). Tento terminační signál musí být lokalizován v *cis* a musí být ve vzdálenosti alespoň 30 nt od terminačního nukleotidu transkripce, ačkoliv krátké oligonukleotidy obsahující signál v *trans* také stimulují terminaci transkripce *in vitro* (HAGLER *et al.* 1994; DENG *et al.* 1996; MOHAMED a NILES 2003a). Krátké oligonukleotidy obsahující terminační signál v *trans* lze způsobem závislým na koncentraci použít k předčasné terminaci transkripce *in vitro* a tím i inhibici syntézy RNA časných genů (MOHAMED a NILES 2003b, 2004). Bylo ukázáno, že sekundární struktury RNA by mohly mít roli v nerozpoznání terminačního signálu transkripce uprostřed kódující sekvence časného genu *A18L* *in vitro* a *in vivo* (PACHA *et al.* 1990; LUO a SHUMAN 1991). Hlubokým sekvenováním transkriptomu VACV bylo zjištěno, že transkripty pouze dvou třetin časných genů obsahují kanonický 5'-UUUUUNU-3' signál v průměrné vzdálenosti 40 nt od terminačního nukleotidu transkripce, což značí existenci dalšího mechanismu tvorby 3' konců časných mRNA či terminace transkripce (YANG *et al.* 2011a). U transkriptů všech časných genů však byl detekován zvýšený obsah pyrimidinových nukleotidů v oblasti pokrývající posledních 25 nt před 3' koncem mRNA (YANG *et al.* 2011a).

### 2.2.6.1 VETF (DE<sub>x</sub>D/H box helikáza)

Zatímco purifikovaná RNAP VACV není schopna syntézy RNA na dvouvláknových DNA templátech (BAROUDY a MOSS 1980; SPENCER *et al.* 1980), rozpustné extrakty purifikovaných částic VACV a méně čisté formy izolované RNAP jsou schopny definovaně iniciovat transkripci časných genů na dvouvláknových DNA templátech *in vitro* (ROHRMANN a MOSS 1985; BROYLES a MOSS 1987b). Později byl identifikován a purifikován časný transkripční faktor viru *Vaccinia* (VETF), který se specificky váže na časnou promotorovou sekvenci UCE *in vitro* a zároveň se neváže na sekvenci UCE s bodovými mutacemi a promotorovou sekvenci pozdních genů (YUEN *et al.* 1987; BROYLES *et al.* 1988). Faktor VETF vykazuje DNA-dependentní ATPázovou aktivitu, jež může být zodpovědná za hydrolýzu  $\beta$ - $\gamma$  vazby ATP, která je nutná pro iniciaci transkripce časných genů VACV (GERSHOWITZ *et al.* 1978; BROYLES a MOSS 1988). Faktor VETF, skládající se z virových proteinů A7 a D6, je spolu s purifikovanou RNAP a DNA templátem obsahujícím časný promotor dostačující pro rekonstituci efektivního transkripčního systému *in vitro* (BROYLES *et al.* 1988; BROYLES a FESLER 1990; GERSHON a MOSS 1990). VETF tvoří nejenom sekvenčně specifické kontakty s UCE sekvencí proti směru transkripce, ale také nespecifické kontakty s nukleotidy +7 až +10 po směru transkripce a ohýbá DNA v promotorové oblasti (BROYLES *et al.* 1991). Poté, co se na komplex DNA:VETF naváže RNA polymeráza (LI a BROYLES 1993b; BALDICK *et al.* 1994), vede ATPázová aktivita VETF k jeho autodisociaci od DNA a následnému zahájení transkripce (BROYLES 1991; LI a BROYLES 1993a). Tyto výsledky vedly k vytvoření modelu iniciace transkripce časných genů VACV (Obrázek 11).

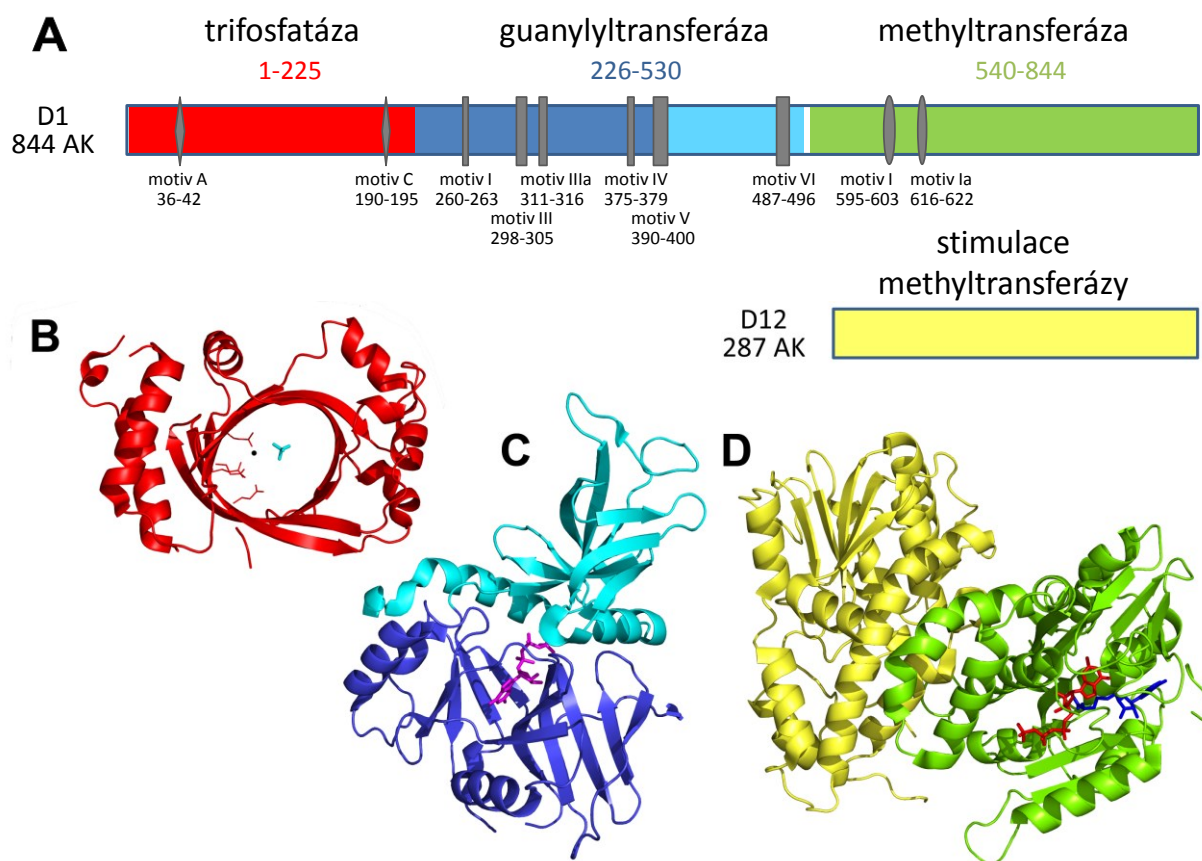


**Obrázek 11:** Model iniciace transkripce časných genů viru *Vaccinia*. (A) Časný transkripční faktor viru *Vaccinia* (VETF, žlutě) se specificky váže na promotorový element (UCE, černě) v průměrné vzdálenosti 12 nt od iniciačního nukleotidu transkripce (šipka, černě), kterým je většinou purinový nukleotid (Pu, šedě), a nespecificky se váže na +7 až +10 místo po směru transkripce (DS, tmavě šedě). (B) Na komplex DNA:VETF se váže RNA polymeráza (RNAP, hnědě), díky interakci N-koncové části asociovaného proteinu (RAP94, oranžově) s VETF. (C) ATPázová aktivita VETF uvolní VETF z preiniciačního komplexu a umožní RNA polymeráze únik z promotoru a syntézu RNA (zeleně). Průměrná vzdálenost iniciačního nukleotidu transkripce od počátku otevřeného čtecího rámce je 22 nt. Upraveno podle (BROYLES 2003).

Za DNA-dependentní ATPázovou aktivitu VETF je zodpovědný menší protein D6, který patří mezi helikázy superrodiny 2 z rodiny DExD/H box helikáz (GORBALENYA *et al.* 1989; LI a BROYLES 1993a). Helikázová aktivita VETF však *in vitro* detekována nebyla (LI a BROYLES 1993a). Obě podjednotky VETF jsou potřebné pro stabilní spojení VETF s DNA promotoru *in vitro* (LI a BROYLES 1993a). Nejasné však zůstává, která podjednotka VETF interaguje s UCE, protože studie využívající *crosslink* si protirečí (BROYLES a LI 1993; CASSETTI a MOSS 1996). Některé substituční mutace aminokyselinových zbytků proteinu D6 narušují ATPázovou a/nebo DNA vazebnou aktivitu VETF (LI a BROYLES 1993a; LI *et al.* 1994; LI a BROYLES 1995). Funkce většího proteinu A7 faktoru VETF je zřejmě závislá na fosforylačním stavu aminokyselinového zbytku Y<sup>367</sup> (NOVY *et al.* 2018). Faktor VETF se může na virovou RNAP vázat pouze v přítomnosti asociovaného proteinu RAP94 *in vitro* a *in vivo*, přičemž tato interakce nevyžaduje žádné další přidatné proteiny (YANG a MOSS 2009).

#### 2.2.6.2 mRNA *capping* enzym (VTF)

*Capping* enzym VACV, multifunkční proteinový komplex složený ze dvou podjednotek, byl prvním purifikovaným a charakterizovaným *capping* enzymem (MARTIN *et al.* 1975; SHUMAN *et al.* 1980). Tento komplex je schopen vykonávat všechny aktivity potřebné pro syntézu mRNA čepičky typu *cap* 0 (MARTIN a MOSS 1975; TUTAS a PAOLETTI 1977; VENKATESAN *et al.* 1980). Podjednotky heterodimerního *capping* enzymu byly identifikovány jako virové proteiny D1 a D12 (MORGAN *et al.* 1984; NILES *et al.* 1989). Aminokyselinové zbytky 1–545 proteinu D1 tvoří autonomní modul s RNA-trifosfatázovou a RNA-guanylyltransferázovou doménou (SHUMAN 1989; SHUMAN a MORHAM 1990; MYETTE a NILES 1996a, b) a aminokyselinové zbytky 548–844 proteinu D1 heterodimerizované s proteinem D12 tvoří autonomní modul vykazující RNA guanin-N<sup>7</sup>-methyltransferázovou aktivitu (CONG a SHUMAN 1992; HIGMAN *et al.* 1992; HIGMAN *et al.* 1994; DE LA PEÑA *et al.* 2007). Guanin-N<sup>7</sup>-methyltransferázová doména se nachází na proteinu D1. Vlastní methyltransferázová aktivita proteinu D1 je však slabá a vyžaduje alosterickou stimulaci proteinem D12 (MAO a SHUMAN 1994; SCHWER *et al.* 2006). Jednotlivé domény a aktivní místa *capping* enzymu VACV byly rozsáhle studovány a charakterizovány pomocí mutagenese a měření enzymových aktivit *in vitro* a *in vivo* (CONG a SHUMAN 1993, 1995; MAO a SHUMAN 1996; YU a SHUMAN 1996; YU *et al.* 1997; HO *et al.* 2000; SAHA a SHUMAN 2001; GONG a SHUMAN 2003; SAHA *et al.* 2003; ZHENG a SHUMAN 2008a, b). Nedávno byla určena krystalová struktura *capping* enzymu viru *Vaccinia* (KYRIELEIS *et al.* 2014), jehož celková organizace je poměrně unikátní (Obrázek 12).



**Obrázek 12:** Schématické zobrazení sekvence a terciární struktury mRNA *capping* enzymu viru *Vaccinia*. (A) 5' RNA trifosfatázová doména (červeně), mRNA-guanylyltransferázová doména (modře a tyrkysově) a mRNA cap (guanin- $N^7$ )-methyltransferázová doména (zeleně) proteinu D1, a doména stimulující methyltransferázu (žlutě) proteinu D12 je zobrazena v odpovídajícím měřítku. Konzervované motivy trifosfatázy (šedé kosočtverce) jsou popsány podle (PEI *et al.* 1999; KYRIELEIS *et al.* 2014). Konzervované motivy guanylyltransferázy (šedé obdélníky) jsou popsány podle (LARSEN *et al.* 1998). Konzervované motivy I a Ia (Ia někdy označován jako II) methyltransferázy (šedé ovály) jsou pojmenovány a popsány podle (WANG a SHUMAN 1997; DE LA PEÑA *et al.* 2007). (B) Terciární struktura D1 trifosfatázy (PDB ID: 4CKB, červeně), pokrývající aminokyselinové zbytky 1–225, s  $Mn^{2+}$  iontem (černě) a analogem  $\gamma$ -fosfátu (tyrkysově). Postranní řetězce čtyř univerzálně konzervovaných kyselých aminokyselinových zbytků motivu A a C, interagujících s dvoumocným kationtem, jsou zobrazeny prostorovým rozložením atomů. Zobrazení ligandů je založeno na struktuře RNA-trifosfatázy Cet1 v komplexu s  $Mn^{2+}$  a sulfátem (PDB ID: 1D8H), s níž byla struktura porovnána pomocí DaliLite v. 3 a vizualizována pomocí PyMOL 1.3. (C) Terciární struktura D1 guanylyltransferázy (PDB ID: 4CKB) v komplexu s GTP (fialově), obsahující nukleotidyltransferázovou doménu (modře) a oligonukleotid vázající doménu (tyrkysově), pokrývající aminokyselinové zbytky 226–528. (D) Terciární struktura D1 methyltransferázy (PDB ID: 4CKB, zeleně), pokrývající aminokyselinové zbytky 548–844, a D12 podjednotky *capping* enzymu (PDB ID: 4CKB, žlutě) v komplexu s S-adenosyl-L-homocysteinem (modře). Struktura 7-methyl-guanosin 5' trifosfátu (červeně) pochází ze struktury mRNA cap (guanin- $N^7$ )-methyltransferázy Ecm1 (PDB ID: 1RI1), s níž byla struktura porovnána a vizualizována pomocí PyMOL 1.3. Poměr mezi relativní velikostí terciárních struktur Obrázku 12 a strukturních modelů Obrázku 3 je zachován.

*Capping* enzym VACV má však kromě syntézy mRNA čepičky typu *cap* 0 ve virovém transkripčním aparátu i další role. Částečně purifikovaná RNA polymeráza VACV za přítomnosti VETF faktoru není schopná terminovat transkripci *in vitro* na templátu obsahujícím časný



terminační signál 5'-TTT<sup>+</sup>TTNT-3', byl však identifikován transkripční terminační faktor viru *Vaccinia* (VTF), komplementující tuto schopnost (SHUMAN *et al.* 1987). Faktor VTF je kopurifikován při izolaci *capping* enzymu VACV, který byl již dříve identifikován jako součást purifikovaného transkripčního komplexu VACV, schopného definované iniciace a terminace transkripce templátů časných genů *in vitro* (BROYLES a MOSS 1987b; SHUMAN *et al.* 1987). *Capping* enzym interaguje s RNA polymerázou VACV *in vitro* i bez přítomnosti nukleových kyselin, a dokáže stimulovat terminaci transkripce po přidání k ternárním elongačním komplexům RNAP *in vitro*, čímž byla identita *capping* enzymu jakožto VTF prokázána (LUO *et al.* 1991; HAGLER a SHUMAN 1992a). Obě podjednotky *capping* enzymu jsou v intaktním stavu esenciální pro terminaci transkripce na terminačním signálu časných genů *in vitro*, ačkoliv bodové mutace, inaktivující trifosfatázovou, guanylyltransferázovou i methytransferázovou aktivitu, v katalytických centrech enzymu nemají na terminaci transkripce *in vitro* vliv (SHUMAN 1989; LUO *et al.* 1995; CONDIT *et al.* 1996a; YU *et al.* 1997). *Capping* enzym je RNA vazebný protein, který vykazuje mírnou preferenci pro U-bohaté sekvence (LUO a SHUMAN 1993). Bylo ukázáno, že interaguje s 5'-UUUUUNU-3' signálem v nascentní RNA při transkripci *in vitro* skrz autonomní modul proteinu D1 s RNA-trifosfatázovou a RNA-guanylyltransferázovou doménou (HAGLER *et al.* 1994; CHRISTEN *et al.* 2008). Pomocí termosenzitivní mutanty proteinu D1 se podařilo prokázat roli *capping* enzymu VACV v terminaci transkripce časných genů *in vivo* (SHATZER *et al.* 2008; TATE *et al.* 2016). *Capping* enzym v přítomnosti terminačního signálu transkripce inhibuje elongaci transkripce tím, že indukuje pauzu elongačního komplexu RNA polymerázy (TATE a GOLLNICK 2015). Efektivita a přesné místo terminace transkripce je však kromě terminačního signálu ovlivněna sekvencí a možná i strukturou nascentní RNA po směru transkripce od terminačního signálu transkripce (TATE a GOLLNICK 2015).

### 2.2.6.3 NPH I (DExD/H box helikáza)

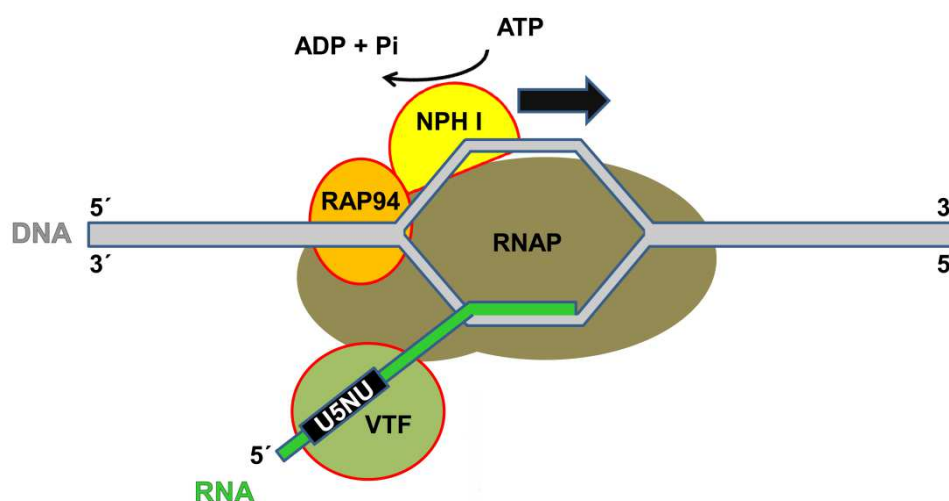
Celkově je efektivita a přesné místo terminace transkripce časných genů VACV určeno kinetickou rovnováhou mezi mírou signalizace terminace transkripce a mírou elongace transkripce, kdy snížení koncentrace substrátů RNA polymerázy VACV vede k posunu míst terminace transkripce blíže k terminačním signálům *in vitro* (HAGLER *et al.* 1994; DENG a SHUMAN 1997). Terminace transkripce časných genů *in vitro* však kromě terminačního signálu vyžaduje ještě hydrolýzu ATP či dATP (SHUMAN *et al.* 1987). Při snížení koncentrace těchto hydrolyzovatelných substrátů dochází k posunu míst terminace transkripce dále od terminačních signálů (HAGLER *et al.* 1994). Předpokládalo se, že za hydrolýzu ATP při terminaci transkripce časných genů je zodpovědný *capping* enzym VACV, protože vykazuje ATPázovou aktivitu

lokalizovanou ve své RNA-trifosfatázové doméně (SHUMAN *et al.* 1980; MYETTE a NILES 1996b). Později však bylo prokázáno, že *capping* enzym VACV není zodpovědný za hydrolýzu ATP během terminace transkripce *in vitro* díky proteinu D1 s mutacemi v této doméně (DENG a SHUMAN 1996; YU *et al.* 1997). Záhy byl pomocí *in vitro* komplementačních analýz identifikován neznámý protein s ATPázovou aktivitou, nutnou pro terminaci transkripce a uvolnění transkriptů z elongačních komplexů, jenž byl nazván nukleosidtrifosfátfosfohydroláza I (NPH I) (DENG a SHUMAN 1998; CHRISTEN *et al.* 1998). NPH I je těsněji spojena s RNA polymerázou VACV než *capping* enzym, který je při purifikacích RNAP uvolněn i nízkou koncentrací soli, zatímco uvolnění NPH I z RNAP vyžaduje heparin (DENG a SHUMAN 1996; CHRISTEN *et al.* 2008).

NPH I je ssDNA-dependentní ATPáza se substrátovou specifitou pro ATP a dATP kódovaná genem *D11L* (PAOLETTI a MOSS 1974; RODRIGUEZ *et al.* 1986; BROYLES a MOSS 1987a; KÜNZI a TRAKTMAN 1989). Podmíněně letální mutace genu tohoto proteinu vede k defektům genové exprese VACV *in vivo* (KÜNZI a TRAKTMAN 1989). NPH I patří mezi helikázy superrodiny 2 z rodiny DExD/H box helikáz (GORBALENYA *et al.* 1989). Pomocí substituční mutagenese NPH I bylo zjištěno, že čtyři helikázové motivy NPH I jsou esenciální pro ATPázovou aktivitu a pět helikázových motivů je esenciálních pro aktivitu NPH I při terminaci transkripce (CHRISTEN *et al.* 1998; MARTINS *et al.* 1999). Helikázová aktivita NPH I na dvouvláknových DNA a RNA templátech však *in vitro* detekována nebyla (DENG a SHUMAN 1998; CHRISTEN *et al.* 1998). Protein NPH I interaguje *in vitro* a *in vivo* svým C-koncem s RNA polymerázou skrz přímou interakci s RAP94 proteinem, který je zároveň esenciální pro terminaci transkripce *in vitro* (MOHAMED a NILES 2000, 2001; PIACENTE *et al.* 2003). NPH I může uvolňovat transkripty z RNAP *in vitro* i nezávisle na *capping* enzymu a terminačním signálu transkripce, pokud je RNA polymeráza zpomalena nebo zastavena nukleotidovou deprivací (PIACENTE *et al.* 2008), a proto je předpokládán dvoukrokový model terminace transkripce časných genů viz Obrázek 13. Předpokládá se, že ATPázová aktivita NPH I je spojena s uvolněním nascentní RNA z elongačních komplexů RNAP mechanismem dopředného posunu RNAP po templátu po směru transkripce bez přítomnosti syntézy RNA, čímž dojde k odstranění 3' konce RNA z aktivního místa RNAP, zkrácení délky DNA:RNA hybridní molekuly a celkové destabilizaci elongačního komplexu (TATE a GOLLNICK 2011). Jako zdroj jednovláknové DNA, sloužící jako kofaktor proteinu NPH I, byla identifikována část netemplátového DNA vlákna transkripční bubliny proti směru transkripce, na které NPH I vykazuje 5' → 3' translokázovou aktivitu (HINDMAN a GOLLNICK 2016).

Kromě role v terminaci transkripce NPH I funguje též jako transkripční elongační faktor, neboť podporuje pročtení oligo(dT) úseků, fungujících jako místa vnitřní pauzy elongačního

komplexu RNAP, za nízkých koncentrací UTP *in vitro* (DENG a SHUMAN 1998). Role NPH I v elongaci transkripce však není esenciální, neboť elongace transkripce *in vitro* probíhá i bez přítomnosti NPH I (CHRISTEN *et al.* 1998).



**Obrázek 13:** Model terminace transkripce časných genů viru *Vaccinia*. Prvním krokem terminace transkripce je rozeznání terminačního signálu transkripce 5'-UUUUUUNU-3' (U5NU, černě) v nascentní RNA (RNA, zeleně) capping enzymem (VTF, olivově), které způsobí pozastavení elongačního komplexu RNA polymerázy (RNAP, hnědě) poblíž místa terminace transkripce. Druhým krokem terminace transkripce je uvolnění transkriptu z pozastaveného elongačního komplexu RNAP způsobem závislým na hydrolýze ATP. Nukleosidtrifosfátfosfohydroláza I (NPH I, žlutě) je součástí elongačního komplexu RNAP díky přímé interakci svého C-konce s N-koncem RNAP-asociovaného proteinu 94 (RAP94, oranžově) a nachází se poblíž transkripční bubliny proti směru transkripce, kde interaguje s netemplátovým vláknem DNA (DNA, šedě). ATPázová aktivita NPH I je spojena s translokací NPH I vzhledem k netemplátovému vláknem DNA v 5' → 3' směru, což způsobí dopředné posunutí RNAP vzhledem k templátu (černá šipka) bez syntézy RNA, vedoucí k uvolnění nascentní RNA a terminaci transkripce. Upraveno podle (HINDMAN a GOLLNICK 2016).

#### 2.2.6.4 NPH II

Roli v terminaci transkripce časných genů *in vivo* má zřejmě také nukleosidtrifosfátfosfohydroláza II (NPH II), což je helikáza z rodiny DExD/H box helikáz, kódovaná genem *I8R* (PAOLETTI a MOSS 1974; SHUMAN 1992). NPH II je DNA či RNA-dependentní NTPáza bez substrátové specifity, vykazující helikázovou aktivitu na dvouvláknových DNA:DNA, DNA:RNA hybridních molekulách a RNA:RNA molekulách *in vitro* (PAOLETTI a MOSS 1974; BAYLISS a SMITH 1996; GROSS a SHUMAN 1996b; TAYLOR *et al.* 2010). Viriony postrádající NPH II produkují časnou mRNA, která je abnormálně dlouhá a zůstává spojena s jádrem virionu (GROSS a SHUMAN 1996a). Protože NPH II vykazuje efektivní helikázovou aktivitu *in vitro* na DNA:RNA hybridních molekulách s DNA vláknem bohatým na puriny, předpokládá se, že NPH II při terminaci transkripce rozplétá DNA:RNA hybridní molekuly, neboť 3' konce RNA časných genů jsou bohaté na pyrimidinové nukleotidy *in vivo* (TAYLOR *et al.* 2010; YANG *et al.* 2011a).



## 2.2.7 Transkripce středních genů viru *Vaccinia*

Transkripce středních a pozdních genů VACV má mnoho společných prvků, které jsou odlišné od transkripce časných genů. K transkripci středních a pozdních genů VACV může docházet až po replikaci virové genomové DNA a je k ní vyžadována syntéza nových molekul RNA polymerázy (VOS a STUNNENBERG 1988; HOODA-DHINGRA *et al.* 1989). Replikace virové genomové DNA je pro transkripci genů těchto tříd nutná zřejmě díky nedostupnosti virové DNA v jádře částice pro vazbu nově syntetizovaných transkripčních faktorů, neboť přítomnost nahého templátu pro transkripci středních genů *in vivo* komplementuje inhibici replikace DNA VACV (KECK *et al.* 1990). Protein RAP94, asociovaný s virionovou formou RNA polymerázy VACV není pro *in vitro* transkripci postreplikativních genů esenciální (WRIGHT a CORONEOS 1995). Transkripce postreplikativních genů *in vitro* nepodléhá terminaci závislé na transkripčním terminačním signálu časných genů, který se často vyskytuje v sekvencích postreplikativních genů (YUEN a MOSS 1987; VOS *et al.* 1991b; CONDIT *et al.* 1996a). Transkripty středních genů se dají detekovat ~100 minut po synchronní infekci HeLa buněk VACV s nejvyšší detekovanou hladinou mRNA kolem 120 minut a předpokládaným poločasem životnosti ~30 minut (BALDICK a MOSS 1993). Byly identifikovány tři střední transkripční faktory viru *Vaccinia* (*Vaccinia virus intermediate transcription factor*; VITF), které jsou spolu s *capping* enzymem esenciální pro iniciaci transkripce středních genů VACV *in vitro* (Tabulka 4). Společným jevem transkripce postreplikativních genů VACV je také to, že se na tomto procesu kromě virových proteinů podílí i některé buněčné proteiny.

Tabulka 4: Střední transkripční faktory viru <i>Vaccinia</i>				
Faktor	Gen	kDa	Funkce	Reference
VITF-1	<i>E4L</i>	30	Iniciace transkripce/ Podjednotka RNAP	ROSALES <i>et al.</i> (1994a)
VITF-2	buněčný	68	Iniciace transkripce	ROSALES <i>et al.</i> (1994b)
VITF-3	<i>A23L, A8R</i>	44, 33	Iniciace transkripce	SANZ a MOSS (1999)
<i>Capping</i> enzym	<i>D1R, D12L</i>	97, 33	Iniciace transkripce	VOS <i>et al.</i> (1991b)
G3BP/p137	buněčný	110, 66	Elongace transkripce?	KATSAFANAS a MOSS (2004)
TBP	buněčný	38	Iniciace transkripce	KNUTSON <i>et al.</i> (2006)
Elongační faktor	<i>J3R</i>	39	Elongace transkripce - Pozitivní vliv	LATNER <i>et al.</i> (2000)
Elongační faktor	<i>G2R</i>	26	Elongace transkripce - Pozitivní vliv	BLACK a CONDIT (1996)
Elongační faktor	<i>H5R</i>	22	Elongace transkripce - Pozitivní vliv	CRESAWN a CONDIT (2007)
Uvolňující faktor	<i>A18R</i>	57	Elongace transkripce - Negativní vliv	XIANG <i>et al.</i> (1998)

Promotory středních genů VACV jsou přibližně 25 nt dlouhé a umístěné proti směru transkripce (MINER *et al.* 1988; HIRSCHMANN *et al.* 1990; KECK *et al.* 1990). Rozsáhlé mutační analýzy čtyř vybraných promotorů středních genů identifikovaly dva promotorové elementy oddělené mezeríkem, které jsou esenciální pro transkripci *in vitro* a *in vivo* (HIRSCHMANN *et al.* 1990; BALDICK *et al.* 1992; KNUTSON *et al.* 2006). Jádrem promotoru (CE) je přibližně 8–11 nukleotidů dlouhý AT bohatý element promotoru vzdálený přibližně 11 nt proti směru transkripce od iniciátorového regionu (INR), což je 5 nukleotidů dlouhý promotorový element na pozici -1 až +4 s konsenzus sekvencí TAAAT (BALDICK *et al.* 1992; BROYLES 2003; KNUTSON *et al.* 2006). Promotory středních a pozdních genů jsou si celkově stran sekvence a umístění vzhledem ke kódující sekvenci genu dost podobné, což demonstruje i fakt, že několika bodovými mutacemi lze proměnit střední promotor v pozdní, a že řada středních promotorů má duální charakter a funguje i jako pozdní promotory (HIRSCHMANN *et al.* 1990; KNUTSON *et al.* 2006; YANG *et al.* 2011b; YANG *et al.* 2013). Novými metodami hlubokého sekvenování transkriptomu VACV byly určeny konsenzus sekvence promotorů středních a pozdních genů viz Obrázek 14 (YANG *et al.* 2011b; YANG *et al.* 2012). Další mutační analýzy odhalily, že pro duální funkci promotoru ve střední i pozdní fázi exprese VACV je důležitý vysoký obsah T bází na pozici -12 až -8, a že hlavní rozdíl mezi středními a pozdními promotory je vzdálenost jejich jádra promotoru od iniciátorového regionu (KNUTSON *et al.* 2006; YANG *et al.* 2013).

**Obrázek 14:** Konsenzus sekvence promotorů středních a pozdních genů viru *Vaccinia* odvozená pomocí programu MEME ze sekvencí rozprostírajících se 50 nt proti směru transkripce a 4 nt po směru transkripce od translačních start kodonů 93 postreprikativních otevřených čtecích rámců, jež byla identifikována na základě celogenomových analýz transkriptů VACV pomocí hlubokého sekvenování (YANG *et al.* 2011b). Jádra promotorů (CE) a iniciátorové regiony (INR) jsou vyznačeny. Konsenzus sekvence duálních promotorů genů střední i pozdní fáze exprese je odvozena z deseti středních promotorů s nejvyšší aktivitou během pozdní fáze exprese a T-bohatý úsek (T), důležitý pro duální aktivitu promotoru, je vyznačen šedě. Šipka udává místo iniciace transkripce. Upraveno podle (MOSS 2013b; YANG *et al.* 2013).

Promotory přibližně poloviny středních a pozdních genů VACV obsahují také dinukleotid GG hned za iniciátorovým elementem TAAAT po směru transkripce. Na tyto promotory se *in vitro* váže buněčný transkripční faktor Yin Yang 1 (YY1) (BROYLES *et al.* 1999; KNUTSON *et al.* 2009b). Nejdřív se předpokládalo, že YY1 bude fungovat jako transkripční aktivátor, ale recentní výsledky ukazují, že YY1 funguje *in vivo* jako transkripční represor genů VACV, jejichž promotory obsahují zmíněný GG dinukleotid (BROYLES *et al.* 1999; KNUTSON *et al.* 2009b).

Iniciačním nukleotidem transkripce postreplikativních genů *in vivo* a *in vitro* je adenosin (BOONE a MOSS 1977; WRIGHT a MOSS 1987; YANG *et al.* 2012). Postreplikativní mRNA mají obecně poměrně krátké 5' nepřekládané oblasti, jejichž počátky jsou zmapovány do blízkosti iniciátorového elementu promotoru, jak bylo detekováno pomocí štěpení S1 endonukleázou (WEIR a MOSS 1984; ROSEL *et al.* 1986; WEIR a MOSS 1987; KECK *et al.* 1990). Metodou prodlužování primeru bylo prokázáno, že 5' konce transkriptů středních a pozdních genů VACV obsahují netemplátový 5' poly(A) leader o délce ~35 nukleotidů, který není kódovaný genomem (BERTHOLET *et al.* 1987; SCHWER *et al.* 1987; BALDICK a MOSS 1993). Bylo ukázáno, že 5' poly(A) leader mRNA vzniká opakovaným klouzáním RNA polymerázy po třech adenosinových zbytcích (deoxythymidinových zbytcích templátového vlákna DNA) v TAAAT iniciátorovém regionu promotoru *in vitro* a *in vivo* (SCHWER a STUNNENBERG 1988; DAVISON a MOSS 1989b). Mezi jednotlivými transkripty daného genu je délka 5' poly(A) leaderu variabilní a dá se *in vitro* ovlivnit koncentrací ATP (SCHWER a STUNNENBERG 1988; AHN a MOSS 1989). Klouzání na AAA sekvenci při iniciaci transkripce je zřejmě obecná vlastnost RNA polymerázy VACV, protože i některé časné transkripty, jejichž promotory obsahují INR element, obsahují 5' poly(A) leader (AHN *et al.* 1990b; INK a PICKUP 1990; YANG *et al.* 2011a). Místa iniciace transkripce jsou u naprosté většiny postreplikativních genů do vzdálenosti 25 nt od start kodónu translace, kdy start kodón je zpravidla přímo součástí TAAAT elementu (YANG *et al.* 2012). Nepřekládané oblasti 5' konců mRNA tedy obvykle tvoří jen 5' poly(A) leader, jehož mediánová délka je 8 nt pro střední a 11 nt pro pozdní transkripty (YANG *et al.* 2012). Přesná biologická role 5' poly(A) leaderu u postreplikativních transkriptů až do nedávna nebyla známa, ale recentně bylo ukázáno, že 5' poly(A) leader poskytuje transkriptům VACV v infikovaných buňkách translační výhodu nad transkripty hostitelské buňky (DHUNGEL *et al.* 2017).

Jak již bylo zmíněno výše, postreplikativní transkripty VACV jsou značně heterogenní v délce svých 3' konců, přičemž délka jednotlivých transkriptů daného genu se může lišit o tisíce párů bází, což znesnadňuje studium mechanismů tvorby 3' konců mRNA (COOPER *et al.* 1981; MAHR a ROBERTS 1984a; WEIR a MOSS 1984). Heterogenita 3' konců mRNA naznačuje, že jakákoliv sekvence potřebná pro terminaci transkripce se musí vyskytovat poměrně často

a musí být poměrně neefektivní, protože se RNA polymeráza s vysokou frekvencí proče za tuto sekvenci (CONDIT a NILES 2002). Kromě transkriptů několika genů, jejichž 3' konce jsou generovány endonukleolytickým štěpením RNA (kapitola 2.2.5.2), nebyl detekován vliv specifické sekvence na terminaci transkripce středních a pozdních genů *in vitro* a *in vivo* (CONDIT a NILES 2002). Pomocí celogenomových analýz transkriptů VACV byl však u postreplikativních transkriptů stejně jako předtím u časných transkriptů detekován zvýšený obsah pyrimidinových nukleotidů v oblasti pokrývající posledních ~25 nt před 3' koncem mRNA (YANG *et al.* 2012).

### 2.2.7.1 VITF-1, VITF-2, VITF-3, mRNA *capping* enzym

Pomocí buněčné frakcionace byly identifikovány čtyři proteinové faktory potřebné pro expresi RNA z promotorů středních genů *A1L*, *A2L*, *I3L* a *G8R* v *in vitro* transkripčním systému odvozeném od extraktů buněk infikovaných VACV (VOS *et al.* 1991a; ROSALES *et al.* 1994a; SANZ a MOSS 1998).

Střední transkripční faktor 1 viru *Vaccinia* (VITF-1) byl identifikován jako podjednotka Rpo30 virové RNA polymerázy se sekvenční podobností s eukaryotickým transkripčním elongačním faktorem TFIIS (ROSALES *et al.* 1994a).

Faktor VITF-2 byl původně identifikován jako blíže neurčený buněčný protein (ROSALES *et al.* 1994b). Při dalších izolacích faktoru VITF-2, využívajících purifikovanou RNAP a ostatní známé transkripční faktory středních genů, byl VITF-2 identifikován jako heterodimer složený z buněčných proteinů G3BP a p137 (KATSAFANAS a MOSS 2004). Proteiny G3BP a p137 jsou dohromady (překvapivě i každý samostatně) esenciální pro transkripci lineárních a otevřených kružnicových templátů *in vitro*, ale zároveň postradatelné pro transkripci superspiralizovaných templátů *in vitro* (KATSAFANAS a MOSS 2004). Díky RNA vazebné doméně proteinu G3BP se předpokládá role obou proteinů v interakci s nascentní RNA VACV během elongace transkripce (KATSAFANAS a MOSS 2004). Oba proteiny se také v buňkách infikovaných VACV přemísťují do virových továren, kde kolokalizují s virovou RNA (KATSAFANAS a MOSS 2007).

Třetí faktor VITF-3 byl identifikován jako heterodimer virových proteinů A8 a A23, které neobsahují žádný rozpoznatelný funkční motiv (SANZ a MOSS 1999). V souladu s předpokládanou rolí faktoru VITF-3 v transkripci VACV lokalizuje protein A23 *in vivo* do virových továren, přičemž oba proteiny VITF-3 faktoru v buňkách asociují s virovou DNA (KATSAFANAS a MOSS 2007; SENKEVICH *et al.* 2017).

Čtvrtý faktor potřebný pro iniciaci transkripce z promotoru středního genu *in vitro* byl identifikován jako *capping* enzym VACV (VOS *et al.* 1991b). Funkce *capping* enzymu v iniciaci transkripce středních genů vyžaduje obě podjednotky, není však závislá na enzymatických

aktivitách tohoto proteinu v syntéze 5' mRNA čepičky na *in vitro* (HARRIS *et al.* 1993; CONDIT *et al.* 1996a). Díky charakterizaci termosenzitivní mutanty proteinu D1 se podařilo prokázat roli *capping* enzymu VACV v transkripci středních genů *in vivo* (SHATZER *et al.* 2008; TATE *et al.* 2016).

#### 2.2.7.2 TATA-box vazebný protein

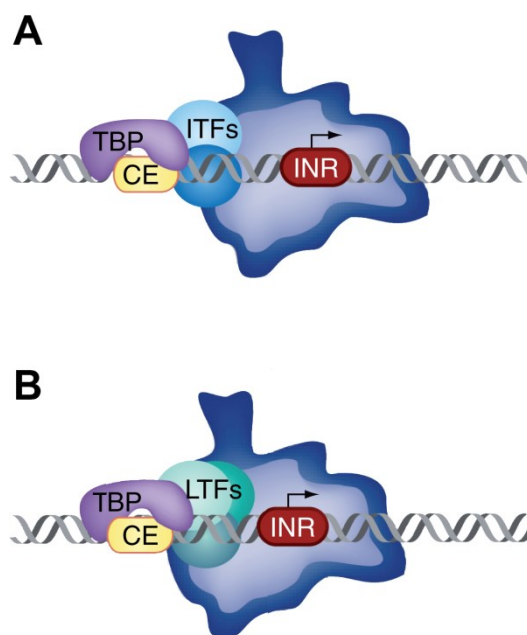
Přestože byly identifikovány proteinové faktory potřebné pro transkripci středních genů VACV, jejich přesné role jsou zatím neznámé. Pokusy s transkripčním systémem *in vitro* dále naznačují, že iniciace transkripce středních genů vyžaduje hydrolýzu  $\beta$ - $\gamma$  vazby ATP, a že k tání promotoru, závislém na ATP, může docházet ještě před vazbou RNA polymerázy k templátu (VOS *et al.* 1991a). U žádného z výše zmíněných faktorů však zatím nebyla popsána ATPázová ani DNA vazebná aktivita *in vitro*. Poměrně nedávno však byla naznačena role DNA vazebné aktivity TATA-box vazebného proteinu (TBP) v transkripci postreplikativních genů VACV.

Protein TBP interaguje *in vitro* s postreplikativními promotory VACV v místě AT bohatého jádra promotoru (KNUTSON *et al.* 2006). Kromě toho protein TBP lokalizuje do virových továren a snížení množství TBP v buňce nebo jeho inaktivace vede ke snížení míry exprese reportérových genů z postreplikativních promotorů *in vivo* (OH a BROYLES 2005; KNUTSON *et al.* 2006). Transkripce postreplikativních genů VACV je inhibována antibiotikem distamycin A, což je ligand malého žlábků DNA, cílící na AT bohaté sekvence (ABU-DAYA *et al.* 1995; BROYLES *et al.* 2004). Protein TBP váže TATA-box v malém žlábků a jeho vazba k DNA je senzitivní na přítomnost distamycinu A (KIM *et al.* 1993a; KIM *et al.* 1993b; CHIANG *et al.* 1994). Předpokládaná interakce TBP s promotory postreplikativních genů VACV tedy vysvětluje citlivost viru k distamycinu A (KNUTSON *et al.* 2006). V aktuálním a obecně přijímaném modelu iniciace transkripce středních a pozdních genů VACV je CE element promotoru vázán TBP proteinem, a střední a pozdní transkripční faktory propojují komplex TBP:DNA s RNA polymerázou, aby došlo ke správnému umístění INR elementu dané třídy promotoru do aktivního místa enzymu viz Obrázek 15 (KNUTSON *et al.* 2006; BROYLES a KNUTSON 2010).

#### 2.2.7.3 Negativní elongační faktor A18

Genetické a biochemické experimenty využívající především termosenzitivní mutanty VACV identifikovaly virové faktory A18, G2, J3 a H5, ovlivňující délku postreplikativních transkriptů *in vivo*, a fungující tedy jako pozitivní a negativní regulátory elongace transkripce (BLACK a CONDIT 1996; XIANG *et al.* 1998; LATNER *et al.* 2000; CRESAWN a CONDIT 2007).





**Obrázek 15:** Model iniciace transkripce středních a pozdních genů viru *Vaccinia*. TATA-box vazebný protein (TBP, fialově) se specificky váže na jádro promotoru (CE, žlutě) středních (A) a pozdních (B) genů. Komplex TBP s DNA (šedě) je cílen středními transkripčními faktory (ITFs, modře) nebo pozdními transkripčními faktory (LTFs, zeleně) a dojde k vytvoření preiniciačního komplexu, na který se váže RNA polymeráza (modrofialově), která iniciuje syntézu RNA na iniciátorovém regionu promotoru (INR, červeně). Díky odlišné vzdálenosti jádra promotoru od iniciátorového regionu je časově regulována exprese jednotlivých tříd genů. Střední transkripční faktory tvoří neaktivní preiniciační komplex na kratším pozdním promotoru a pozdní transkripční faktory tvoří neaktivní preiniciační komplex na delším středním promotoru. K přepnutí střední transkripce na pozdní transkripci dojde akumulací pozdních transkripčních faktorů, které z preiniciačních komplexů vytěsní střední transkripční faktory. Upraveno podle (BROYLES a KNUTSON 2010).

Virové mutanty proteinu A18 vykazují aberantní postreplikativní transkripční profil známý jako promiskuitní transkripce, při které dochází k transkripci virového genomu, který běžně při pozdní fázi životního cyklu viru přepisován není, což vede k celkovému zvýšení množství dvouvláknové RNA v infikovaných buňkách a aktivaci protivirové dráhy využívající RNázu L (BAYLISS a CONDIT 1993). Další analýzy zmíněné mutanty ukázaly, že v infikovaných buňkách dochází k syntéze transkriptů středních genů, které jsou delší než obvykle, takže protein A18 má vlastnosti negativního transkripčního elongačního faktoru (XIANG *et al.* 1998). Protein A18 je DNA-dependentní ATPáza se substrátovou specifitou pro ATP a dATP, patří do rodiny DExD/H box helikáz (KOONIN a SENKEVICH 1992; BAYLISS a CONDIT 1995). ATPázová aktivita proteinu A18 je *in vitro* stimulována jednovláknovou a dvouvláknovou DNA a také DNA:RNA hybridní molekulou (BAYLISS a CONDIT 1995). Protein A18 též vykazuje helikázovou aktivitu *in vitro*, schopnou rozplétat krátké dvouvláknové molekuly DNA o délce do 20 nt ve 3' → 5' směru (SIMPSON a CONDIT 1995). Další pokusy ukázaly, že protein A18 je esenciální pro uvolnění nascentní RNA z elongačního komplexu RNAP VACV *in vitro* způsobem závislým na ATP (LACKNER a CONDIT 2000). Protein A18 je faktor uvolňující transkripty z elongačního komplexu RNAP, který ale pro svoji uvolňující aktivitu *in vitro* vyžaduje přítomnost blíže neurčeného buněčného faktoru (LACKNER a CONDIT 2000).

Pro identifikaci pozitivních elongačních faktorů postreplikativní transkripce VACV se osvědčila především izolace a charakterizace mutant VACV rezistentních k nebo závislých



na léčivu zvaném isatin- $\beta$ -thiosemikarbazon (IBT), což je chemická sloučenina s antipoxvirovým účinkem strukturně připomínající purinový nukleosid (SHEFFIELD *et al.* 1960). Infikované buňky vystavené léčbě pomocí IBT vykazují stejný aberantní transkripční profil postreplikativních genů VACV jako termosenzitivní mutanty proteinu A18 a vyznačují se syntézou delších postreplikativních transkriptů (BAYLISS a CONDIT 1993).

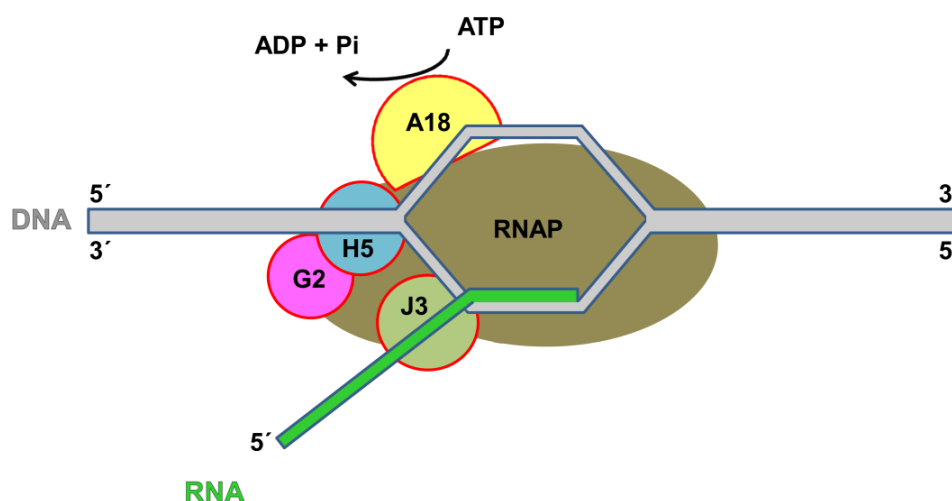
#### 2.2.7.4 Pozitivní elongační faktory G2, J3 a H5

Role proteinu G2 v elongaci transkripce VACV byla původně odhalena mapováním mutací přinášejících růstovou závislost VACV na IBT (MEIS a CONDIT 1991). Transkripty středních a pozdních genů VACV jsou v kmenech s mutací genu *G2R* v nepřítomnosti IBT na 3' koncích zkrácené oproti transkriptům divokého typu, což je opačný fenotyp, než vykazují mutanty genu *A18R* (BLACK a CONDIT 1996). Faktor G2 tedy nějakým způsobem pozitivně ovlivňuje elongaci transkripce VACV *in vivo*. Vliv proteinů A18 a G2 na regulaci stejného procesu byl demonstrován vzájemným genetickým vztahem, kdy mutace genu *G2R* funguje jako supresor mutace genu *A18R* (CONDIT *et al.* 1996b). Protein G2 nemá žádné další známé aktivity, ale bylo ukázáno, že interaguje s proteinem H5, což je multifunkční protein, fungující například jako stimulační pozdní transkripční faktor VACV a zřejmě i jako pozitivní elongační faktor transkripce postreplikativních genů VACV (KOVACS a MOSS 1996; BLACK *et al.* 1998; MCCRAITH *et al.* 2000; CRESAWN a CONDIT 2007).

Izolace a charakterizace dalších mutant VACV, růstově závislých na IBT, a mutant, fungujících jako supresory mutací genu *A18R*, identifikovala produkt genu *J3R* VACV jako druhý faktor pozitivně ovlivňující elongaci transkripce VACV *in vivo* (LATNER *et al.* 2000). Mutanty genu *J3R* vykazují obdobný fenotyp jako mutanty *G2R* genu, tj. zkrácení transkriptů středních a pozdních genů VACV na 3' koncích v nepřítomnosti IBT (XIANG *et al.* 2000). Role proteinu J3 v elongaci transkripce postreplikativních genů VACV je svým způsobem překvapivá, protože protein J3 je stimulatorní podjednotkou virové poly(A) polymerázy a zároveň virovou mRNA *cap* (nukleosid-2'-O)-methyltransferázou (GERSHON *et al.* 1991; SCHNIERLE *et al.* 1992). Role proteinu J3 v elongaci postreplikativní transkripce VACV je však na těchto aktivitách nezávislá *in vivo* (XIANG *et al.* 2000; LATNER *et al.* 2002).

Role produktu virového genu *H5R* v elongaci transkripce postreplikativních genů VACV byla předpokládána především díky asociaci proteinu H5 s elongačním faktorem G2 (BLACK *et al.* 1998). Pomocí cílené mutagenese genu *H5R* však byla získána mutanta VACV rezistentní k IBT, což silně naznačuje roli H5 jako pozitivního elongačního faktoru transkripce VACV (CRESAWN a CONDIT 2007).

Předpokládá se, že mutanty VACV rezistentní k IBT mají intermediární fenotyp pohybující se mezi citlivostí a závislostí na IBT, kdy v nepřítomnosti IBT je většina postreplikativních transkriptů VACV dostatečně dlouhá, aby kódovala funkční proteiny, a v přítomnosti IBT není většina postreplikativních transkriptů VACV tak dlouhá, aby mohlo vzniknout významné množství dvouvláknové RNA (LATNER *et al.* 2002; PRINS *et al.* 2004). Byly také identifikovány a charakterizovány další mutanty VACV rezistentní k IBT, které obsahují bodové mutace v genech *J6R* a *A24R* kódujících dvě největší podjednotky RNA polymerázy VACV (CONDIT *et al.* 1991; PRINS *et al.* 2004; CRESAWN *et al.* 2007). Komplexy RNAP se zmíněnými bodovými mutacemi vykazují defekty v elongaci transkripce *in vitro*, což naznačuje, že RNAP je v centru kontroly elongace transkripce, která je zprostředkována dalšími faktory (PRINS *et al.* 2004; CRESAWN *et al.* 2007). Přestože cíl a přesný mechanismus účinku IBT na elongaci transkripce VACV není znám, podařilo se alespoň částečně charakterizovat složení transkripčního elongačního a terminačního komplexu postreplikativních genů (Obrázek 16).



**Obrázek 16:** Model transkripčního elongačního a terminačního komplexu středních a pozdních genů viru *Vaccinia*. Transkripty postreplikativních genů viru mají 3' konce heterogenní délky. Elongace transkripce je ovlivňována několika virovými faktory. Faktor H5 (H5, tyrkysově) interaguje s proteinem G2, a jedná se o multifunkční protein, který funguje jako pozdní transkripční faktor, faktor štěpící 3' konce mRNA některých postreplikativních genů a pravděpodobně jako pozitivní elongační faktor, podporující komplex RNA polymerázy (RNAP, hnědě) v elongaci. Protein G2 (G2, fialově) interaguje s proteinem H5 a chová se jako pozitivní elongační faktor transkripce *in vivo*. Faktor J3 (J3, olivově) funguje jako mRNA *cap* 1 methyltransferáza, stimulatívni podjednotka poly(A) polymerázy a pozitivní elongační faktor transkripce *in vivo*. Faktor J3 je zřejmě i součástí transkripčních elongačních komplexů časných genů, protože interaguje s proteinem RAP94 a NPH I. Protein A18 (A18, žlutě) je DNA helikáza, která se chová jako negativní transkripční elongační faktor *in vivo*. ATPázová aktivita A18 je spojena s uvolněním nascentní RNA (RNA, zeleně) z transkripčního komplexu *in vitro*, za pomoci dalšího zatím neznámého buněčného faktoru. Nejpravděpodobnější místo akce proteinu A18 je úsek netemplátového vlákna DNA (DNA, šedě) transkripční bubliny proti směru transkripce, jenž se nachází na povrchu transkripčního komplexu RNA polymerázy, a po kterém může A18 translokovat. Upraveno podle (CONDIT a NILES 2002).

## 2.2.8 Transkripce pozdních genů viru *Vaccinia*

Stejně jako transkripce středních genů vyžaduje transkripce pozdních genů viru předchozí replikaci virové genomové DNA, nicméně pouhá přítomnost nahého templátu pro transkripci pozdních genů viru *in vivo* za podmínek inhibice replikace DNA VACV nestačí (VOS a STUNNENBERG 1988; KECK *et al.* 1990). Transkripce pozdních genů viru je, kromě replikace virové genomové DNA, závislá ještě na *de novo* proteinové syntéze, neboť pro expresi pozdních genů *in vivo* z nahého templátu za podmínek inhibice replikace DNA VACV je potřeba *de novo* syntéza tří pozdních transkripčních faktorů VACV (VLTF), které jsou kódovány středními geny viru (VOS a STUNNENBERG 1988; KECK *et al.* 1990). Celkem bylo identifikováno několik stimulačních či esenciálních virových a buněčných faktorů, nezbytných pro transkripci pozdních genů *in vivo* a *in vitro* (Tabulka 5). Transkripty pozdních genů se dají detekovat ~140 minut po synchronní infekci HeLa buněk VACV se vzrůstající hladinou mRNA do 6 hodin po infekci a jejich předpokládaný poločas životnosti je ~30 minut (BALDICK a MOSS 1993).

Tabulka 5: Pozdní transkripční faktory viru <i>Vaccinia</i>				
Faktor	Gen	kDa	Funkce	Reference
VLTF-1	<i>G8R</i>	30	Iniciace transkripce	KECK <i>et al.</i> (1990)
VLTF-2	<i>A1L</i>	17	Iniciace transkripce	KECK <i>et al.</i> (1990)
VLTF-3	<i>A2L</i>	26	Iniciace transkripce	KECK <i>et al.</i> (1990)
VLTF-4	<i>H5R</i>	22	Iniciace transkripce?	KOVACS a MOSS (1996)
VLTF-X	buněčný	35-40	Iniciace transkripce	WRIGHT <i>et al.</i> (1998)
hnRNP A2/ RBM3	buněčný	35, 20	Iniciace transkripce? Elongace transkripce?	WRIGHT <i>et al.</i> (2001)
TBP	buněčný	38	Iniciace transkripce	KNUTSON <i>et al.</i> (2006)
Elongační faktor	<i>J3R</i>	39	Elongace transkripce - Pozitivní vliv	LATNER <i>et al.</i> (2000)
Elongační faktor	<i>G2R</i>	26	Elongace transkripce - Pozitivní vliv	BLACK a CONDIT (1996)
Elongační faktor	<i>H5R</i>	22	Elongace transkripce - Pozitivní vliv	CRESAWN a CONDIT (2007)
Uvolňující faktor	<i>A18R</i>	57	Elongace transkripce - Negativní vliv	XIANG <i>et al.</i> (1998)

Původní deleční analýzy ukázaly, že promotory pozdních genů VACV jsou krátké, umístěné do vzdálenosti 20 nt proti směru transkripce před předpokládaným místem iniciace transkripce, a poukázaly také na důležitost konzervovaného iniciátorového elementu promotoru TAAAT (BERTHOLET *et al.* 1986; HÄNGGI *et al.* 1986; WEIR a MOSS 1987). Rozsáhlé mutační analýzy dvou vybraných pozdních promotorových oblastí identifikovaly dva promotorové elementy oddělené mezerníkem, které jsou esenciální pro transkripci *in vivo* (DAVISON a MOSS 1989b; KNUTSON *et al.* 2006). Jádrem promotoru je přibližně 9–11 nukleotidů dlouhý AT bohatý

element promotoru vzdálený přibližně 6 nt proti směru transkripce od iniciátorového regionu (DAVISON a MOSS 1989b; BROYLES 2003; KNUTSON *et al.* 2009a). V rámci iniciátorového regionu dochází k iniciaci transkripce a syntéze 5' poly(A) leaderu mRNA *in vitro* a *in vivo*, přičemž mutace uvnitř a poblíž INR elementu mohou délku 5' poly(A) leaderu mRNA ovlivnit jak *in vitro*, tak *in vivo* (WRIGHT a MOSS 1987; DE MAGISTRIS a STUNNENBERG 1988; SCHWER a STUNNENBERG 1988; DAVISON a MOSS 1989b). Pro sílu pozdního promotoru je důležitý obsah T nukleotidů jádra promotoru, kdy syntetické promotory obsahující 18–20 po sobě jdoucích T nukleotidů dosahují vyšší míry exprese, než přirozeně se vyskytující promotory (DAVISON a MOSS 1989b). Výše zmíněnou strukturní a funkční podobnost promotorů středních a pozdních genů VACV demonstruje i fakt, že stejné jádro promotoru může současně fungovat jako střední a pozdní promotor v dvousměrných promotorech (KNUTSON *et al.* 2009a), a mutacemi ovlivňujícími délku mezerníku mezi CE a INR lze změnit specifitu středního promotoru na pozdní a *vice versa* (YANG *et al.* 2013). Stejně jako v případě středních genů VACV nebyl detekován vliv specifické sekvence na terminaci transkripce pozdních genů *in vitro* a *in vivo* (CONDIT a NILES 2002), ačkoliv elongace a terminace transkripce pozdních genů viru je ovlivňována stejnými proteinovými faktory, jako v případě transkripce středních genů viru (kapitoly 2.2.7.3 a 2.2.7.4).

#### 2.2.8.1 VLTF-1, VLTF-2, VLTF-3, VLTF-4

Původní pokusy s *in vitro* transkripčním systémem odvozeným od buněčných extraktů infikovaných buněk odhalily, že iniciace transkripce pozdních genů VACV nevyžaduje hydrolýzu  $\beta$ - $\gamma$  vazby ATP (SCHWER a STUNNENBERG 1988). Pomocí buněčné frakcionace byl částečně purifikován první pozdní transkripční faktor viru *Vaccinia* (*Vaccinia virus late transcription factor*; VLTF) nazvaný VLTF-1 (WRIGHT a MOSS 1989), a následně byly identifikovány virové geny tři transkripčních faktorů, které jsou esenciální a dostatečné pro transaktivaci exprese reportérového genu z pozdního promotoru VACV *in vivo* v infikovaných buňkách za podmínek inhibice replikace DNA VACV (KECK *et al.* 1990). Jedná se o geny *A1L*, *A2L* a *G8R*, které jsou přepisovány ve střední fázi životního cyklu VACV, díky čemuž byl navržen kaskádový model regulace genové exprese VACV (KECK *et al.* 1990).

Dříve purifikovaný transkripční faktor VLTF-1 byl následně identifikován jako virový protein G8, a jeho esenciální role při transkripci pozdních genů VACV byla prokázána *in vitro* a *in vivo* (WRIGHT *et al.* 1991; ZHANG *et al.* 1992). Protein G8 interaguje sám se sebou a také s proteiny A1, A2 a H5 (MCCRAITH *et al.* 2000; DELLIS *et al.* 2004). Predikce sekundární a terciární struktury proteinu G8 naznačují, že by se mohlo jednat o trimerní protein tvořící

DNA svorku, orthologní eukaryotickému *proliferating cell nuclear antigen* proteinu, což je procesivní faktor replikace jaderné DNA (DA SILVA a UPTON 2009).

Jako VLTF-2 byl pojmenován virový protein A1, který je taktéž esenciální pro transkripci pozdních genů VACV *in vitro* a *in vivo* (KECK *et al.* 1990; KECK *et al.* 1993b; WRIGHT a CORONEOS 1993). Termosenzitivní virové mutanty proteinu A1 vykazují v nepermisivních podmínkách fenotyp, při kterém nedochází k syntéze proteinů pozdních genů viru (CARPENTER a DELANGE 1992). Protein A1 interaguje kromě proteinu G8 také sám se sebou a může tvořit homodimery *in vitro* (DELLIS *et al.* 2004). Protein A1 je konzervovaný mezi všemi nukleocytoplasmatickými viry a obsahuje předpokládaný motiv zinkového prstu (KECK *et al.* 1993a; IYER *et al.* 2006).

Jako VLTF-3 byl pojmenován protein A2, který také obsahuje předpokládaný motiv zinkového prstu a dokáže *in vitro* vázat zinek (KECK *et al.* 1993a). Bodové mutace předpokládaného motivu zinkového prstu snižující afinitu proteinu k zinku *in vitro* brání proteinu A2 v transaktivaci exprese z pozdního virového promotoru *in vivo* (KECK *et al.* 1993a). Esenciální role VLTF-3 při transkripci pozdních genů VACV byla také později prokázána v *in vitro* transkripčním systému (HUBBS a WRIGHT 1996; PASSARELLI *et al.* 1996). Protein A2 interaguje kromě proteinu G8 také s proteinem H5 (DELLIS *et al.* 2004). Faktory VLTF-1, 2 a 3 asociují s virovou DNA *in vivo*, a předpokládá se, že tvoří proteinový komplex (SENKEVICH *et al.* 2017).

Multifunkční virový protein H5, který stimuluje transkripci pozdních genů *in vitro* přibližně 40násobně, byl později pojmenován jako VLTF-4, (KOVACS *et al.* 1994; KOVACS a MOSS 1996). Role tohoto proteinu v transkripci pozdních genů VACV nebyla odhalena v dřívějších *in vivo* komplementačních studiích, neboť protein H5 je exprimován i v časně fázi životního cyklu viru, tedy ještě před replikací virové DNA (KOVACS *et al.* 1994). Protein H5 interaguje s některými pozdními transkripčními faktory a postreplikativními elongačními faktory transkripce (BLACK *et al.* 1998; DELLIS *et al.* 2004). Kromě toho interaguje protein H5 s virovou proteinkinázou B1 a virovým replikačním proteinem A20 (MCCRAITH *et al.* 2000; ISHII a MOSS 2002). Protein H5 může být fosforylovaný na dvou threoninových zbytcích proteinkinázou B1 *in vitro* a *in vivo* a virovou proteinkinázou F10 *in vitro* (BEAUD *et al.* 1995; TRAKTMAN *et al.* 1995; BEAUD a BEAUD 2000; BROWN *et al.* 2000). Jak již bylo zmíněno výše (kapitola 2.2.5.2 a 2.2.7.4), protein H5 je s největší pravděpodobností také postreplikativní transkripční elongační faktor a štěpící faktor generující 3' konce některých postreplikativních transkriptů endonukleolytickým štěpením (CRESAWN a CONDIT 2007; D'COSTA *et al.* 2008). Kromě transkripce virových genů má protein H5 ještě roli v morfogenezi virových částic a replikaci genomové DNA VACV *in vivo* (DEMASI a TRAKTMAN 2000; D'COSTA *et al.* 2010; BOYLE *et al.* 2015). Biochemické a biofyzikální vlastnosti proteinu H5 naznačují, že má nestrukturovaný N-konec,  $\alpha$ -helikální

C-konec a primárně se vyskytuje v tetramerní formě, přičemž jeho fosforylace způsobuje agregaci do multimerních komplexů vyššího řádu (KAY *et al.* 2013). Protein H5 vykazuje endonukleolytickou aktivitu na jednovláknových RNA substrátech generující 3'-OH konce a váže preferenčně dvouvláknové nukleové kyseliny způsobem nezávislým na jejich sekvenci *in vitro* (KAY *et al.* 2013).

#### 2.2.8.2 VLTF-X, hnRNP A2/RBM3, TATA-box vazebný protein

Další pokusy využívající buněčnou frakcionaci identifikovaly pátý transkripční faktor VLTF-X, esenciální pro *in vitro* transkripční systém specifický pro pozdní geny VACV (WRIGHT a CORONEOS 1993). Ačkoliv se předpokládalo, že jde o virový protein, bylo ukázáno, že se jedná o buněčný protein, nacházející se v infikovaných i neinfikovaných buňkách v cytoplasmatické i jaderné frakci (GUNASINGHE *et al.* 1998; WRIGHT *et al.* 1998). Faktor VLTF-X má DNA vazebnou aktivitu a dokáže specificky interagovat se sekvencemi pozdních promotorů VACV *in vitro* (GUNASINGHE *et al.* 1998; WRIGHT *et al.* 1998; ZHU *et al.* 1998). Vazba faktoru VLTF-X k pozdnímu promotoru *in vitro* není závislá na přítomnosti ATP a  $Mg^{2+}$  iontů (GUNASINGHE *et al.* 1998). Jeho vazba k pozdnímu promotoru *in vitro* není ovlivněna mutacemi v TAAAT iniciátorovém elementu promotoru, nicméně v případě delece AT bohatého jádra promotoru již k vazbě faktoru VLTF-X nedochází (GUNASINGHE *et al.* 1998; ZHU *et al.* 1998).

Snaha o identifikaci faktoru VLTF-X pomocí jiného purifikačního postupu vedla k izolaci buněčného heterogenního jaderného ribonukleoproteinu A2 (hnRNP A2) a buněčného proteinu s RNA vazebným motivem 3 (RBM3) (WRIGHT *et al.* 2001). Oba tyto proteiny nezávisle na sobě stimulují transkripci pozdních genů viru *in vitro* ze superspiralizovaných, ale ne z lineárních templátů (WRIGHT *et al.* 2001). Oba tyto proteiny dokážou také *in vitro* vázat jednovláknovou a dvouvláknovou DNA, ačkoliv zmíněné interakce postrádají dříve detekovanou sekvenční specifitu pro pozdní promotor (WRIGHT *et al.* 2001). Role proteinů hnRNP A2 a RBM3 by mohla být v iniciaci či elongaci transkripce VACV (WRIGHT *et al.* 2001). Protein hnRNP A2 byl také nalezen asociovaný s virovou DNA *in vivo* (SENKEVICH *et al.* 2017).

Je všeobecně uznáváno, že původně izolovaný faktor VLTF-X odpovídá proteinům hnRNP A2 a RBM3 (BROYLES 2003; BROYLES a KNUTSON 2010; MOSS 2013b). Domnívám se však, že původně izolovaný faktor VLTF-X mohl být TATA-box vazebný protein, což by bylo v souladu s molekulovou hmotností faktoru VLTF-X a DNA vazebnými vlastnostmi faktoru VLTF-X, specifickými pro sekvence pozdních promotorů (GUNASINGHE *et al.* 1998; ZHU *et al.* 1998). Protein TBP (na rozdíl od proteinů hnRNP A2 a RBM3) interaguje s postreplikativními promotory VACV *in vitro* specificky (WRIGHT *et al.* 2001; KNUTSON *et al.* 2006).



### 3 CÍLE PRÁCE

Hlavním cílem této práce bylo charakterizovat transkripční aparát lineárních plasmidů kvasinky *Kluyveromyces lactis* bioinformatickými, genetickými a biochemickými metodami.

Bioinformatická charakterizace *in silico*:

- Určit fylogenetický původ nekanonické RNA polymerázy a plasmidových promotorů.
- Odhalit sekvenční elementy zodpovědné za netemplátovou 5' polyadenylaci některých plasmidových transkriptů.
- Odhalit sekvenční nebo strukturní elementy podílející se na tvorbě 3' konců plasmidových transkriptů.
- Určit identitu plasmidových otevřených čtecích rámců s dosud neznámou funkcí.

Genetická a biochemická charakterizace:

- Ověřit nezbytnost methyltransferázové domény mRNA *capping* enzymu pro fungování plasmidů *in vivo*.
- Ověřit předpokládané interakce mezi všemi složkami transkripčního aparátu *in vivo*, případně detekovat nové interakční partnery purifikovaného transkripčního komplexu.
- Otestovat RNA polymerázovou aktivitu purifikovaného transkripčního komplexu *in vitro*.
- Otestovat interakci předpokládané helikázy s plasmidovou DNA *in vivo* a *in vitro*.
- Otestovat vliv nalezeného sekvenčního elementu promotoru na 5' polyadenylaci plasmidových transkriptů *in vivo*.
- Otestovat vliv nalezených strukturních elementů na tvorbu 3' konců plasmidových transkriptů *in vivo*.

Dalšími cíli této práce, v návaznosti na určení předpokládaného původu transkripčního aparátu lineárních plasmidů, byla charakterizace nepřekládaných oblastí transkriptů viru *Vaccinia* z čeledi *Poxviridae* bioinformatickými, genetickými a biochemickými metodami.

Bioinformatická, genetická a biochemická charakterizace:

- Otestovat přítomnost mRNA čepičky na vybraných poxvirových transkriptech *in vivo*.
- Detekovat 3' konce vybraných postreplicativních transkriptů poxvirů *in vivo*.
- Identifikovat sekvenční nebo strukturní elementy podílející se na tvorbě 3' konců postreplicativních transkriptů poxvirů.

## 4 MATERIÁL A METODY

### 4.1 MATERIÁL

#### 4.1.1 Chemikálie

Název	Výrobce	Vzorec/Zkratka
[ $\alpha$ – $^{32}\text{P}$ ] UTP	MGP	
[ $\gamma$ – $^{32}\text{P}$ ] ATP	MGP	
8-hydroxychinolin	Lachema	
$\beta$ -merkaptoethanol	Sigma	
adenin	ANGUS	
agar	ForMedium™	
agarosa	Serva / Invitrogen	
akrylamid	Serva / BIO-RAD	
ampicilin	Spofa / Sigma	
anti-HA agarosa, HA-7	Sigma	
azid sodný	Advondale Laboratories Limited	
B-PER® bakteriální lyzační roztok	Thermo Fisher Scientific	
bayol F	Serva	
bovinní sérum albumin	Serva	BSA
bromfenolová modř	Merck	
<i>Coomassie Brilliant Blue G-250</i>	Serva	
deionizovaná voda		ddH <sub>2</sub> O
deoxyribonukleosidtrifosfáty (samostatně i ve směsi)	Fermentas / Roche	dNTP
destilovaná voda		dH <sub>2</sub> O
diethylpyrokarbonát	Sigma	DEPC
dihydrogenfosforečnan draselný	Lachema	KH <sub>2</sub> PO <sub>4</sub>
dimethylsulfoxid	Sigma	DMSO
dithiotreitol	AMRESCO	DTT
dodecylsírán sodný	Serva	SDS
dusičnan stříbrný	Lachema	AgNO <sub>3</sub>
dusík kapalný	Linde Gas	
EcoLume™ Liquid Scintillation Cocktail	ICN	
ethanol	Lach-Ner	
ethidium bromid	Sigma	EtBr
fenol	Merck	
fenol saturovaný kyselý, pH 4,5	MP Biomedicals	

fenylmethylsulfonyl fluorid	Sigma	PMSF
formaldehyd	Sigma	
formamid	Sigma	
G418	Amersham / Sigma	
galaktosa	Sigma / Serva	
GFP-Trap®_A	ChromoTek	
glukosa	Serva	
Glutathione Sepharose™ 4 Fast Flow	Amersham	
glutaraldehyd	Serva	
glutathion	Sigma	
glycerol	Sigma / Lach-Ner	
glycin	Serva	
GST-Trap®_A	ChromoTek	
hydrogenfosforečnan disodný	Lachema	Na <sub>2</sub> HPO <sub>4</sub>
hydrogenuhličitan sodný	Lachema	NaHCO <sub>3</sub>
hydroxid sodný	Lachema	NaOH
hygromycin B Gold™	InvivoGen	HygB
chloramfenikol	Lepetit	
chlorid cesný	ANGUS	CsCl
chlorid draselný	Lach-Ner	KCl
chlorid hořečnatý	Sigma	MgCl <sub>2</sub>
chlorid manganatý	Sigma	MnCl <sub>2</sub>
chlorid sodný	Lach-Ner	NaCl
chlorid zinečnatý	Lachema / Sigma	ZnCl <sub>2</sub>
chloroform	Lach-Ner	
imidazol	Sigma	
isoamylalkohol	Lachema	
isopropanol	Lach-Ner	
isopropyl-β-D-thiogalaktosid	Fermentas / Sigma	IPTG
jednovláknová DNA fága M13	NEB	
jednovláknová DNA z lososího mlíčí	Sigma / Roche	
kanamycin	Serva	
kvasničný autolyzát	Imuna / ForMedium™	
kyselina citrónová	Lachema	
kyselina ethylendiamintetraoctová	Serva	EDTA
kyselina chlorovodíková	Lachema / ICN	HCl
kyselina p-kumarová	Sigma	
kyselina octová	Lach-Ner	
kyselina octová ledová	Lach-Ner	
kyselina trichloroctová	Sigma	TCA
L-histidin	Serva	

L-leucin	ANGUS	
L-lysin	Nutritional Biochemicals Co.	
L-methionin	Merck	
L-tryptofan	Reanal	
LightShift™ Poly (dI-dC)	Thermo Fisher Scientific	
lineární polyakrylamid GenElute™	Sigma	
luminol	Sigma	
methanol	Lach-Ner / Lachema	
minerální olej	Sigma	
N,N'-methylenbisakrylamid	Serva / Lachema	
N,N,N',N'-tetramethylethylendiamin	Sigma	TEMED
náhodné hexanukleotidové primery	Invitrogen	
Ni-NTA Superflow	Quiagen	
Nodidet P-40	Sigma	
octan draselný	Sigma	KAc
octan litný	Sigma	LiAc
octan sodný	Sigma	NaAc
pepton	OXOID/ ForMedium™	
peroxid vodíku	Sigma / Lach-Ner	
persíran amonný	Serva	APS
polyethylenglykol 4000	Serva	
protein G-agarosa	Sigma	
pyrofosforečnan sodný	Sigma	
ribonukleosidtrifosfáty (samostatně i ve směsi)	Thermo Fisher Scientific	NTP
RNA Blue	Top-Bio	
sacharosa	Serva	
síran amonný	Lach-Ner	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>
směs inhibitorů proteáz cOmplete™ Mini	Roche	
sorbitol	Serva / Sigma	
Sunar Complex 1	Hero	
SYBR™ Gold	Molecular Probes	
TRI Reagent® Solution	Ambion	
tris(hydroxymethyl)-aminomethan	Serva / ANGUS	Tris
Triton X-100	Serva	
Tween 20	Serva / Sigma	
uhličitan sodný	Lachema	NaCO <sub>3</sub>
uracil	Sigma	
X-Gal	Sigma	
xylenová modř	Merck	
<i>Yeast Nitrogen Base</i>	ANGUS / Difco / Sigma	

### 4.1.2 Roztoky a pufrы

**1x PBS:** 140 mM NaCl; 2,7 mM KCl; 10 mM Na<sub>2</sub>HPO<sub>4</sub>; 1,8 mM KH<sub>2</sub>PO<sub>4</sub>; pH 7,3

**1x TBS:** 50 mM Tris-HCl pH 7,5; 150 mM NaCl

**4x Tris-HCl/SDS pH 6,8:** 0,5 M Tris-HCl; 0,4 % (w/v) SDS

**4x Tris-HCl/SDS pH 8,8:** 1,5 M Tris-HCl; 0,4 % (w/v) SDS

**5x SDS-PAGE pufr:** 25 mM Tris; 192 mM glycín; 0,1 % (w/v) SDS

**6x vzorkový pufr pro agarosovou elektroforézu:** 10 mM Tris-HCl pH 7,6;  
0,03 % (w/v) bromfenolová modř; 60 % (v/v) glycerol; 60 mM EDTA-NaOH

**6x vzorkový pufr pro SDS-PAGE:** 53,8 % (v/v) 4x Tris-HCl/SDS pH 6,8;  
38,5 % (v/v) glycerol; 7,6 % (v/v) β-merkaptóethanol; 270 mM SDS;  
0,57 mM bromfenolová modř

**10x TBE pufr:** 0,9 M Tris; 0,9 M kyselina boritá; 0,02 M EDTA-NaOH pH 8,0

**10x transkripční pufr:** 400 mM Tris-HCl pH 8; 30 mM MgCl<sub>2</sub>; 10 mM DTT

**50x TAE pufr:** 24,2 % (w/v) Tris; 5,7 % (v/v) kyselina octová ledová;  
10 % (v/v) 0,5 M EDTA-NaOH pH 8,0

**akrylamid + N,N'-methylenbisakrylamid:** 30 % (w/v) akrylamid;  
0,8 % (w/v) N,N'-methylenbisakrylamid; přefiltrovat

**blotovací roztok:** 20 % (v/v) methanol; 0,29 % (w/v) glycín; 0,58 % (w/v) Tris;  
0,037 % (w/v) SDS

**blokovací roztok:** 1x TBS; 0,5 % (v/v) Tween 20; 5 % (w/v) Sunar Complex 1

**EDTA-NaOH:** 0,5 M roztok EDTA upravený pomocí NaOH na pH 8

**fixační roztok:** 25 % (v/v) izopropanol; 10 % (v/v) kyselina octová

**izolační roztok:** 50 mM NaAc pH 5,2; 10 mM EDTA-NaOH pH 8; 1 % (w/v) SDS

**odmývací roztok:** 1x TBS; 0,5 % Tween 20

**PEG:** 50 % (w/v) polyethylenglykol 4000

**převrstvovací pufr pH 8,8:** 0,375 M Tris-HCl; 0,1 % (w/v) SDS

**roztok I (izolace plasmidové DNA):** 25 mM Tris-HCl pH 8; 10 mM EDTA-NaOH;  
50 mM glukosa

**roztok II (izolace plasmidové DNA):** 1 % (w/v) SDS; 0,2 M NaOH

**roztok III (izolace plasmidové DNA):** 3 M KAc; 11,5 % (v/v) kyselina octová

**roztok Coomassie Blue:** 10 % (v/v) kyselina octová;  
0,006 % (w/v) Coomassie Brilliant Blue G-250

**roztok PMSF:** 100 mM PMSF v izopropanolu

**roztok kyseliny p-kumarové:** 93,5 mM kyselina p-kumarová v DMSO

**roztok luminolu:** 250 mM luminol v DMSO

**roztok X-Gal:** 20 mg/ml X-Gal v DMSO

**STET:** 10 % (v/v) sacharosa; 50 mM Tris-HCl pH 8; 50 mM EDTA-NaOH;  
1 % (v/v) Triton X-100

**STOP pufr:** 95 % (v/v) formamid; 0,05 % (w/v) bromfenolová modř; 20 mM EDTA-NaOH; 0,05 % (w/v) xlenová modř

**TE:** 10 mM Tris-HCl pH 8; 1 mM EDTA-NaOH pH 8

**TESor:** 1,2 M sorbitol; 10 mM Tris-HCl pH 8; 25 mM EDTA-NaOH pH 8; 1 % (v/v)  $\beta$ -merkaptoethanol

**TESP:** 20 mM Tris-HCl pH 8; 50 mM EDTA-NaOH pH 8; 2 % (w/v) SDS; před použitím přidat 0,5 mg/ml pronáza E

**TGZD:** 20 mM Tris-HCl pH 8,0; 75 mM KCl; 10  $\mu$ M ZnCl<sub>2</sub>; 5 % (v/v) glycerol

**TKCl:** 0,8 M KCl; 10 mM Tris-HCl pH 7,5

**TKG:** 44 mM Tris-HCl pH 8,0; 30 mM KCl; 2 % (v/v) glycerol

**TM:** 10 mM Tris-HCl pH 7,5; 10 mM MgCl<sub>2</sub>

**TN:** 50 mM Tris-HCl pH 7,5; 150 mM NaCl

**Tris-HCl:** roztok Tris upravený na požadované pH pomocí HCl

**TRN:** 110 mg/ml ribonukleáza A; 0,1 mM Tris-HCl pH 7,5; 0,045 mM NaCl

**TSor:** 1,2 M sorbitol; 10 mM Tris-HCl pH 7,5

**upravený fenol:** vodou nasycený fenol ekvilibrovaný 50 mM Tris-HCl na pH 8; 0,1 % (w/v) 8-hydroxychinolin; 0,1 % (v/v)  $\beta$ -merkaptoethanol

**upravený chloroform:** chloroform + isoamylalkohol v poměru 24:1

#### 4.1.2.1 Pufry pro purifikaci proteinů a imunoprecipitaci

##### 4.1.2.1.1 Purifikace proteinů s GST-epitopem

**omývací pufr:** TKG

**eluční pufr 1:** 50 mM Tris-HCl pH 8; 10 mM glutathion pH 8

**eluční pufr 2:** 50 mM Tris-HCl pH 8; 20 mM glutathion pH 8

##### 4.1.2.1.2 Purifikace proteinů s HIS-epitopem

**omývací pufr:** TN + 10 mM imidazol; 2 mM  $\beta$ -merkaptoethanol

**eluční pufr 1:** TN + 100 mM imidazol; 2 mM  $\beta$ -merkaptoethanol

**eluční pufr 2:** TN + 250 mM imidazol; 2 mM  $\beta$ -merkaptoethanol

##### 4.1.2.1.3 Purifikace proteinů s GFP-epitopem pomocí GFP-Trap<sup>®</sup>\_A

**GFP wash:** 10 mM Tris-HCl pH 7,5; 150 mM NaCl; 0,5 mM EDTA-NaOH; těsně před použitím přidat směs inhibitorů proteáz (1 tableta směsi inhibitorů cOmplete<sup>™</sup> Mini na 10 ml) a 1 mM PMSF



#### 4.1.2.1.4 Imunoprecipitace a chromatinová imunoprecipitace

**NDLB (nedenaturující lyzační pufr):** 1 % (v/v) Triton X-100; 50 mM Tris-HCl pH 7,4; 300 mM NaCl; 5 mM EDTA-NaOH; 0,02 % (w/v) azid sodný; těsně před použitím přidat inhibitory proteáz (1 tableta směsi cOmplete<sup>TM</sup> Mini na 10 ml) a 1 mM PMSF

**ChIP wash:** 20 mM Tris-HCl pH 8,0; 150 mM NaCl; 2 mM EDTA-NaOH;

1 % (v/v) Triton X-100; 0,1 % (w/v) SDS

**ChIP final wash:** ChIP wash + 500 mM NaCl

**eluční pufr:** 1 % (w/v) SDS; 100 mM NaHCO<sub>3</sub>

#### 4.1.3 Komerční pufry\*

**pufr pro deoxyribonukleázu I:** 10 mM Tris-HCl pH 7,5 při 25 °C; 2,5 mM MgCl<sub>2</sub>;

0,5 mM CaCl<sub>2</sub> (Ambion / Invitrogen)

**pufr pro poly(A) polymerázu:** 50 mM Tris-HCl pH 7,9 při 25 °C; 250 mM NaCl (Applied Biosystems)

**pufr pro restrikční endonukleázu *Bam*HI:** 10 mM Tris-HCl pH 8 při 37 °C; 5 mM MgCl<sub>2</sub>;

100 mM KCl; 0,02 % (v/v) Triton X-100; 0,1 mg/ml BSA (Fermentas)

**pufr pro restrikční endonukleázu *Eco*RI:** 50 mM Tris-HCl pH 7,5 při 37 °C; 10 mM MgCl<sub>2</sub>;

100 mM NaCl; 0,02 % (v/v) Triton X-100; 0,1 mg/ml BSA (Fermentas)

**pufr pro restrikční endonukleázu *Orange*:** 50 mM Tris-HCl pH 7,5 při 37 °C;

100 mM NaCl; 10 mM MgCl<sub>2</sub>; 0,1 mg/ml BSA (Fermentas)

**pufr pro restrikční endonukleázu *Red*:** 10 mM Tris-HCl pH 8,5 při 37 °C; 100 mM KCl;

10 mM MgCl<sub>2</sub>; 0,1 mg/ml BSA (Fermentas)

**pufr pro restrikční endonukleázu *Yellow*:** 33 mM Tris-octan pH 7,9 při 37 °C;

66 mM KAc; 10 mM octan hořečnatý; 0,1 mg/ml BSA (Fermentas)

**pufr pro SAP fosfatázu:** 10 mM Tris-HCl pH 7,5 při 37 °C; 10 mM MgCl<sub>2</sub>; 0,1 mg/ml BSA (Fermentas)

**pufr pro SuperScript III<sup>TM</sup> reverzní transkriptázu:** 50 mM Tris-HCl pH 8,3 při 20 °C;

75 mM KCl; 3 mM MgCl<sub>2</sub> (Invitrogen)

**pufr pro T4 DNA ligázu:** 40 mM Tris-HCl pH 7,8 při 25 °C; 10 mM MgCl<sub>2</sub>; 10 mM DTT;

0,5 mM ATP (Fermentas)

**pufr pro T4 polynukleotid kinázu (forward):** 50 mM Tris-HCl pH 7,6 při 25 °C;

10 mM MgCl<sub>2</sub>; 5 mM DTT; 0,1 mM spermidin (Fermentas / TaKaRa)

**pufr pro Taq DNA polymerázu:** 10 mM Tris-HCl pH 8,3 při 20 °C; 50 mM KCl;

1,5 mM MgCl<sub>2</sub> (Roche)

**pufr pro terminální deoxyribonukleotidyltransferázu (TdT):** 25 mM Tris-HCl pH 6,6

při 25 °C; 200 mM kakodylát draselný; 0,25 mg/ml BSA (Fermentas / Roche)

\* U jednotlivých složek pufrů jsou uvedeny finální koncentrace v reakční směsi.

#### 4.1.4 Enzymy a jiné proteiny

Název	Zásobní koncentrace	Výrobce
benzonáza	25 U/μl	Novagen
deoxyribonukleáza I	2 U/μl	Ambion / Invitrogen
lysozym	500 U/mg	Serva
lytikáza	2000 U/mg	Sigma
Pfu DNA polymeráza	2,5 U/μl	Fermentas / NEB / Roche
Pwo DNA polymeráza	2,5 U/μl	Roche
pronáza E	6 U/mg	Serva / Roche
proteináza K	16,4 mg/μl	Roche / Merck
poly(A) polymeráza	2 U/μl	Applied Biosystems
restrikční endonukleázy	10 U/μl	Fermentas / NEB
ribonukleáza A	58 U/mg	Sigma
RiboLock™	40 U/μl	Fermentas
ribonukleázové inhibitory	20 U/μl	Invitrogen
<i>Shrimp alkaline phosphatase</i> (SAP)	1 U/μl	Fermentas
SuperScript III™ reverzní transkriptáza	200 U/μl	Invitrogen
T4 DNA ligáza	5 Weiss U/μl	Fermentas
T4 polynukleotid kináza	10 U/μl	Fermentas / TaKaRa
Taq DNA polymeráza	1 U/μl; 5 U/μl	Roche / Fermentas
Taq DNA polymeráza , FastStart™	5 U/μl	Roche
TdT	400 U/μl	Roche / Fermentas

#### 4.1.5 Komerční soupravy

**BigDye® Terminator v3.1 Cycle Sequencing Kit** (Applied Biosystems)

**CloneJET PCR Cloning Kit** (Fermentas)

**DNA Clean & Concentrator™-5** (Zymo Research)

**DNA-free™ Kit** (Ambion)

**FavorPrep™ GEL/PCR Purification Kit** (Favorgen)

**FavorPrep™ Plasmid Extraction Mini Kit** (Favorgen)

**FastBack DNA Minispin Kit** (Renogen Biolab)

**FastStart™ PCR Master** (Roche)

**FirstChoice® RLM-RACE Kit** (Invitrogen)

**GenBond Plasmid FlexSpin Kit** (Renogen Biolab)

**High Pure PCR Product Purification Kit** (Roche)

**LightCycler® 480 SYBR Green I Master** (Roche)

**Poly(A) Tailing Kit** (Applied Biosystems)

**QIAquick Nucleotide Removal Kit** (Quiagen)

**TOPO® TA Cloning® Kit** (Invitrogen)

#### 4.1.6 Standardy molekulových hmotností

Všechny použité standardy molekulových hmotností jsou zobrazeny níže na Obrázku 17.

**GeneRuler™ 50 bp Plus DNA Ladder** (Fermentas)

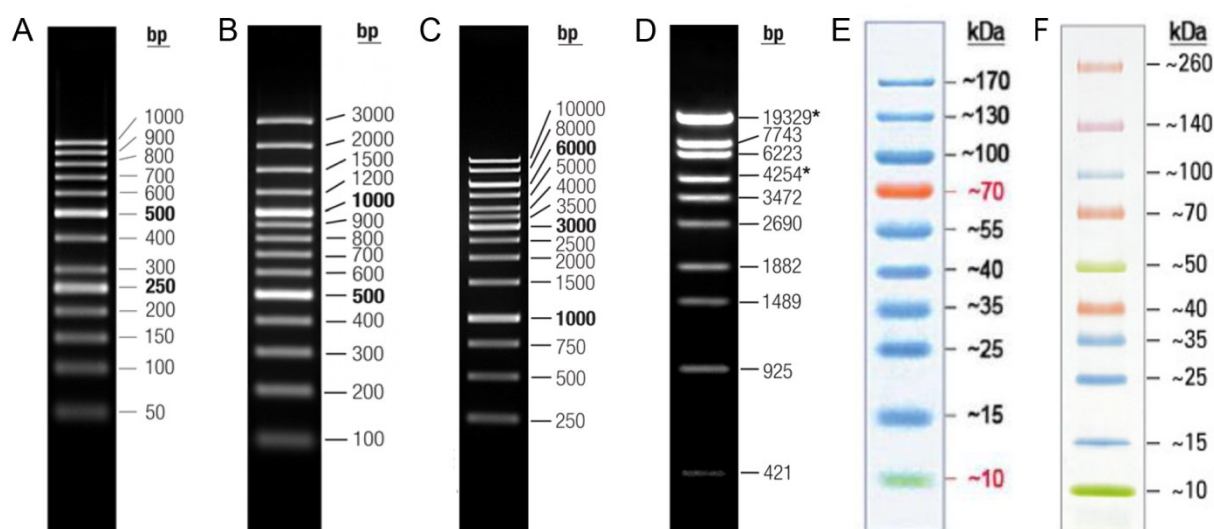
**GeneRuler™ 100 bp Plus DNA Ladder** (Fermentas)

**GeneRuler™ 1 kb DNA Ladder** (Thermo Fisher Scientific)

**Lambda DNA/ *Eco130I* (*StyI*) Marker, 16** (Fermentas)

**PageRuler™ Prestained Protein Ladder** (Fermentas)

**Spectra™ Multicolor Broad Range Protein Ladder** (Thermo Fisher Scientific)



**Obrázek 17:** Standardy molekulových hmotností. Čísla udávají velikost daných fragmentů DNA nebo daných proteinů. Hvězdička označuje fragmenty, které se mohou spojit v jediný fragment tvořící další nezobrazený pruh. (A) GeneRuler™ 50 bp Plus DNA Ladder, (B) GeneRuler™ 100 bp Plus DNA Ladder, (C) GeneRuler™ 1 kb DNA Ladder, (D) Lambda DNA/*Eco130I* (*StyI*) Marker, 16, (E) PageRuler™ Prestained Protein Ladder, (F) Spectra™ Multicolor Broad Range Protein Ladder. Převzato z příbalových letáků daných produktů.

## 4.1.7 Plasmidy a vektory

**Tabulka 6:** Plasmidy a vektory použité v této práci

Název vektoru	Velikost (bp)	Popis relevantních charakteristik	Zdroj / Původ
pcDNA™ 5/FRT/TO	5137	β-laktamáza, počátek replikace pUC, CMV promotor/tetracyclinový operátor, FRT, gen rezistence k hygromycinu B	Invitrogen
pCR™ 4-TOPO®	3956	β-laktamáza, aph(3')-II, počátek replikace pUC, lacZα-ccdB pod kontrolou Lac promotoru, T7 a T3 promotor	Invitrogen
pET16b/His-ORF6 ORF7	9132	β-laktamáza, počátek replikace pBR322, lacIq, Rop, bicistronní His-K2ORF6 K2ORF7 pod kontrolou T7 promotoru, T7 terminátor	Tato práce
pET16b/ORF6	8712	β-laktamáza, počátek replikace pBR322, lacIq, Rop, His-K2ORF6 pod kontrolou T7 promotoru, T7 terminátor	VOPÁLENSKÝ (1999)
pGEX4T2/GST-ORF6 ORF7	8308	β-laktamáza, počátek replikace pBR322, lacIq, lacZα pod kontrolou Lac promotoru, bicistronní GST_TEV-K2ORF6 K2ORF7 pod kontrolou Tac promotoru	Tato práce
pGEX4T2/GST-ORF6 ORF7-His	8338	β-laktamáza, počátek replikace pBR322, lacIq, lacZα pod kontrolou Lac promotoru, bicistronní GST_TEV-K2ORF6 K2ORF7-His pod kontrolou Tac promotoru	Tato práce
pGEX4T2/GST-ORF6	7887	β-laktamáza, počátek replikace pBR322, lacIq, lacZα pod kontrolou Lac promotoru, GST_TEV-K2ORF6 pod kontrolou Tac promotoru	SÝKORA (2013)
pGEX4T2/GST-ORF7	5391	β-laktamáza, počátek replikace pBR322, lacIq, lacZα pod kontrolou Lac promotoru, GST_TEV-K2ORF7 pod kontrolou Tac promotoru	SÝKORA (2013)
pGKL1	8874	plasmid divokého typu, killer +	GUNGE <i>et al.</i> (1981)
pGKL2	13457	plasmid divokého typu	GUNGE <i>et al.</i> (1981)
pJET1.2/blunt	2974	β-laktamáza, počátek replikace pMB1, <i>Eco47I</i> pod kontrolou $P_{lacUV5}$ promotoru, T7 promotor	Fermentas
pRKL1-1	6425	modifikovaný pGKL1 obsahující gen rezistence ke G418 pod kontrolou UCR2 z plasmidu pGKL1 (K1UCR2), deletovaný ORF2 gen, killer -	V. Vopálenský
pRKL1-2	6425	modifikovaný pGKL1 obsahující gen rezistence ke G418 pod kontrolou K1UCR2 s jednou bodovou mutací v předpokládaném iniciátorovém regionu, deletovaný ORF2 gen, killer -	Tato práce
pRKL1-3	6425	modifikovaný pGKL1 obsahující gen rezistence ke G418 pod kontrolou K1UCR2 s dvěma bodovými mutacemi v předpokládaném iniciátorovém regionu, deletovaný ORF2 gen, killer -	Tato práce
pRKL1-4	9970	modifikovaný pGKL1 obsahující gen ORF4-HA pod kontrolou K1UCR4 a gen rezistence k hygromycinu B pod kontrolou K2UCR6	Tato práce
pRKL1-5	6595	modifikovaný pGKL1 obsahující gen rezistence ke G418 pod kontrolou K1UCR2, následovaný 3' nepřekládanou oblastí genu <i>K2ORF5</i> , deletovaný ORF2 gen, killer -	Tato práce
pRKL1-6	6595	modifikovaný pGKL1 obsahující gen rezistence ke G418 pod kontrolou K1UCR2, následovaný 3' nepřekládanou oblastí genu <i>K2ORF5</i> s mutacemi v předpokládané RNA vlásence 2, deletovaný ORF2 gen, killer -	Tato práce
pRKL1-7	6595	modifikovaný pGKL1 obsahující gen rezistence ke G418 pod kontrolou K1UCR2, následovaný 3' nepřekládanou oblastí genu <i>K2ORF5</i> s mutacemi obnavujícími předpokládanou RNA vlásenku 2, deletovaný ORF2 gen, killer -	Tato práce
pRKL1-8	6484	modifikovaný pGKL1 obsahující gen rezistence ke G418 pod kontrolou K1UCR2, následovaný předpokládanou RNA vlásenkou genu A27L viru <i>Vaccinia</i> , deletovaný ORF2 gen, killer -	Tato práce
pRKL1-9	9988	modifikovaný pGKL1 obsahující gen ORF4 pod kontrolou K1UCR4 s dvěma bodovými mutacemi v předpokládaném iniciátorovém regionu a gen rezistence k hygromycinu B pod kontrolou K2UCR6	Tato práce
pRKL1-10	9999	modifikovaný pGKL1 obsahující gen HA-ORF3 pod kontrolou K1UCR3, gen rezistence k hygromycinu B pod kontrolou K2UCR6 a gen ORF4 pod kontrolou K2UCR4	Tato práce

pRKL2-1	14353	modifikovaný pGKL2 obsahující gen ORF2 pod kontrolou K1UCR1 a gen rezistence ke G418 pod kontrolou K1UCR2	Tato práce
pRKL2-2	14353	modifikovaný pGKL2 obsahující bodové mutace D432A, G438A a D440A předpokládaného SAM místa ORF3 genu, gen ORF2 pod kontrolou K1UCR1 a gen rezistence ke G418 pod kontrolou K1UCR2	Tato práce
pRKL2-3	14368	modifikovaný pGKL2 obsahující gen Flag-ORF6 pod kontrolou K2UCR6 a gen rezistence ke G418 pod kontrolou K1UCR2	SÝKORA (2013)
pRKL2-4	15082	modifikovaný pGKL2 obsahující fúzní gen yEGFP3-ORF6 pod kontrolou K2UCR6 a gen rezistence ke G418 pod kontrolou K1UCR2	SÝKORA (2013)
pRKL2-5	16199	modifikovaný pRKL2-4 obsahující gen ORF7-Flag pod kontrolou K2UCR7 a gen rezistence k hygromycinu B pod kontrolou K1UCR1	Tato práce
pRKL2-6	16206	modifikovaný pRKL2-4 obsahující gen ORF3-HA pod kontrolou K2UCR3, gen ORF2 pod kontrolou K1UCR1 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-7	16213	modifikovaný pRKL2-4 obsahující gen HA-ORF4 pod kontrolou K2UCR4 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-8	15104	modifikovaný pGKL2 obsahující fúzní gen yEGFP3-ORF4 pod kontrolou K2UCR4 a gen rezistence ke G418 pod kontrolou K1UCR2	Tato práce
pRKL2-9	16194	modifikovaný pRKL2-8 obsahující gen HA-ORF6 pod kontrolou K2UCR6 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-10	16228	modifikovaný pRKL2-8 obsahující gen ORF3-HA pod kontrolou K2UCR3, gen ORF2 pod kontrolou K1UCR1 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-11	15094	modifikovaný pGKL2 obsahující fúzní gen ORF3-yEGFP3 pod kontrolou K2UCR3, gen ORF2 pod kontrolou K1UCR1 a gen rezistence ke G418 pod kontrolou K1UCR2	Tato práce
pRKL2-12	14547	modifikovaný pGKL2 obsahující gen HA-ORF6 pod kontrolou K2UCR6 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-13	14588	modifikovaný pGKL2 obsahující gen HA-ORF4 pod kontrolou K2UCR4 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-14	14581	modifikovaný pGKL2 obsahující gen ORF3-HA pod kontrolou K2UCR3, gen ORF2 pod kontrolou K1UCR1 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-15	14574	modifikovaný pGKL2 obsahující gen ORF7-Flag pod kontrolou K2UCR7 a gen rezistence k hygromycinu B pod kontrolou K1UCR1	Tato práce
pRKL2-16	16202	modifikovaný pRKL2-4 obsahující gen ORF7-c-Myc pod kontrolou K2UCR7 a gen rezistence k hygromycinu B pod kontrolou K1UCR1	Tato práce
pRKL2-17	14455	modifikovaný pGKL2 obsahující gen D13L viru <i>Vaccinia</i> pod kontrolou K2UCR9, gen ORF10 pod kontrolou K1UCR1, gen rezistence ke G418 pod kontrolou K1UCR2 a deletovaný gen ORF9	Tato práce
pRKL2-18	14793	modifikovaný pGKL2 obsahující fúzní gen ORF6/7 spojený linkerem z genu J6R viru <i>Vaccinia</i> pod kontrolou K2UCR6, gen rezistence ke G418 pod kontrolou K1UCR2 a deletovaný gen ORF6 a ORF7	Tato práce
pRKL2-19	14355	modifikovaný pGKL2 obsahující rozdělený gen ORF6 na dva geny odpovídající $\beta$ a $\beta'$ podjednotkám RNAP pod kontrolami K2UCR6, gen rezistence ke G418 pod kontrolou K1UCR2	Tato práce
pRKL2-20	14611	modifikovaný pGKL2 obsahující gen HA-ORF9 pod kontrolou K2UCR4 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-21	14604	modifikovaný pGKL2 obsahující gen HA-ORF9 pod kontrolou K2UCR4, gen ORF8 pod kontrolou K1UCR1 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-22	16168	modifikovaný pRKL2-4 obsahující gen ORF9-HA pod kontrolou K2UCR4, gen ORF10 pod kontrolou K1UCR1 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-23	16229	modifikovaný pRKL2-4 obsahující gen HA-ORF9 pod kontrolou K2UCR4, gen ORF8 pod kontrolou K1UCR1 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-24	16241	modifikovaný pRKL2-11 obsahující gen HA-ORF9 pod kontrolou K2UCR4, gen ORF8 pod kontrolou K1UCR1 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-25	14538	modifikovaný pGKL2 obsahující gen His-ORF6 pod kontrolou K2UCR6 a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-26	15015	modifikovaný pGKL2 obsahující fúzní gen GST_TEV-ORF6 pod kontrolou K2UCR6 a gen rezistence ke G418 pod kontrolou K1UCR2	Tato práce
pRKL2-27	15359	modifikovaný pGKL2 obsahující gen rezistence ke G418 pod kontrolou promotoru genu D12L viru <i>Vaccinia</i> a gen rezistence k hygromycinu B pod kontrolou K1UCR3	Tato práce
pRKL2-28	15086	modifikovaný pGKL2 obsahující fúzní gen ORF4-yEGFP3 pod kontrolou K2UCR4, gen ORF5 pod kontrolou K1UCR1 a gen rezistence ke G418 pod kontrolou K1UCR2	Tato práce
pUG36	6225	CEN6/ARSH4, URA3, $\beta$ -laktamáza, počátek replikace colE1, yEGFP3 pod kontrolou MET25 promotoru, CYC1 terminátor	J. H. Hegemann, GenBank: AF298791.1

pUG6	4009	$\beta$ -laktamáza, počátek replikace colE1, loxP-gen rezistence ke G418-loxP	GÜLDENER <i>et al.</i> (1996)
------	------	---	-------------------------------

### 4.1.8 Oligonukleotidové primery

**Tabulka 7:** Oligonukleotidové primery použité v této práci

Název primeru	Nukleotidová sekvence 5' → 3'
3RACE_VV_A11R	AGGACTGAGTACCGATATCAC
3RACE_VV_A21L	GTTCCATTCCATGTTCCAAG
3RACE_VV_A26L	TCATCGGACACTATTGTTGTG
3RACE_VV_A27L	AACTCTCAAACAACGGCTAAC
3RACE_VV_A28L	ATAACGCAGAGTGTCAGTTTC
3RACE_VV_A45R	CAATTAACGAGAATGGCGTTTG
3RACE_VV_D10R	AGGGGTGAGAATGTTCCAGAG
3RACE_VV_D6R	GACGCTAGGAATTACATAGATG
3RACE_VV_H3L	TACAGGAACATTCTGTACCGC
3RACE_VV_I5L	ACGCTACGTACGTCAAATCCC
3RACE_VV_I8R	CTCGCTGAGGAGTTGGATAAC
3RACE_VV_O2L	CCCGAAAATGAATTGCGTGAC
5' RACE Adapter	GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA
5' RACE Inner Primer	CGCGGATCCGAACACTGCGTTTGTGCTTGTATG
5' RACE Outer Primer	GCTGATGGCGATGAATGAACACTG
5_RACE_O2_K1	CCTTGACTCCATAATTTGCAGCT
5_RACE_O4_K1	TCCATTAAATCCAGAGTATTCTTTC
5_RACE_O4_K1	TCCATTAAATCCAGAGTATTCTTTC
5_RACE_O8_K2	TCTTTTCAAACATCTAGCCACC
5RACE_O2_K2	TTCGTATGTAAATGTTCCGCA
5RACE_O3_K1_2	GTATCTTCTCGCAAAAAGATCTTC
5RACE_O3_K1_3	TAGGATACCAAATTCCTGAAGGC
5RACE_O4_K2	ATCTAGAATCAAGAACAACCTTCTCA
5RACE_O6_K2	CTGACCAATTTAATGGTAAATTC
5RACE_O7_K2	CAAATAGCTCATTTTGTCTAAGC
5RACE_O10_K2	TTGTTTATATGTGCCAGCA
5RACE_VV_A17L	CTTCTATAGTCTGTCTTTCG
5RACE_VV_A1L	CGTAAACGCCGCTTTATCTC
5RACE_VV_A5R	AGAACCCTCCTCTATCTCTTG
5RACE_VV_C3L	CCTCGGGATTCCATACCATAG
5RACE_VV_D12L	CACGTGGAAGGTAAACATCTT
5RACE_VV_G8R	GAATGACGGTTCTACCACAAC
5RACE_VV_H5R	TTACCAGCTTCAACTGTACC
5RACE_VV_I4L	GGTCTTTCAACGATCTTGTTG
5RACE_VV_J6R	GGTGCATACATTCACTGTTC
actin_KL-RNAfor	TGGTATGTGTAAGCCGGTT
aktin_KL-rev	AACACCGTCACCAGAATCCAA
anch2	GATTGAGGTGTATCTGATGTCGA
c-myc_forw	GAACAAAAATTAATAAGTGAAGAAGATTTA
EGFP-ORF4_1F	GATTATAACAATCAGATTAATAATGAAAGTATTCACTAAGACCGTGAAAGCTACTTAATATAG
EGFP-ORF4_2F	GTTTCATTTGATGCTCGATGAGTTTTTCTAACTATCTGATTATAACAATCAGATTAATAATGTCTAAAGGTGAAGAATTATTC ACTGGTG
EGFP-ORF4_2R	GTATTGTTCTATACTTAAATTTTATTAGTGAATACTTTCATCTGCAGCCCGGGGGATCCACTAGTT
EGFPReverse	TTACTTGACAGCTCGTCCATG
FLAG-rev	TCTTTATCATCATCATCTTTATAATC
HGT1_KL-forw	GTTGCGTTTTGATATCGCATC



HGT1_KL-rev	TGACAACCGTAACCGATGTAG
Hygro_rev	CGGAGACGCTGTGAACT
Hygro_tail	ACGGCAATTTTCGATGATGCAG
in_Kan_rev1	GCAGTGGTGAGTAACCATGCA
in_ORF3_forw	CTCTTGAAATAGAGTTTG
in_ORF5_forw	AGTGGTGAAGAGGAAAAATC
in_ORF5_rev	GAGTAGTCTTTTCCGTATCCT
in_ORF6_rev	TCTCCTTTCTTTCATAGGTGC
in_ORF9_rev	TGGAAATCTATTATGTAAAC
inORF6_forv2	AGGTCTTGAGCATGGAATACT
K05	AAAATGTGTTTAGAATTAGA
K07	AATGAAGATATATCATATATTTAG
K08	CAACTCTGTATAACAAGTCT
K1_HA-ORF3_F2	TTTAAATAAAATGAAGATATATCATATTATATGATCGTATTAGATATACCTAAAATGAAAAAGCCTGAACTACCGCGAC
K1_HA-ORF3_R1	ATATGATATATCTTCATTTTAAATAAAAATGTATCCATATGATGTTCCAGATTATGCTTGTTTGAATTAGATAATAATC
K1_HA-ORF3_R2	GATAACAAACACTAAATATATGATATATTTTCATTTTAAATCTGATTGTTATAATCAGATAGTTATTCCTTTGCCCTCGGACGA GTG
K1_K2O5-UTR_2F	CATTTGATGCTCGATGAGTTTTTCTAAGATTATGGACAAATAGAAATTT
K1_K2O5-UTR_2R	GTTTTCTCTAAAGAAGCAAATACTGGTAAGAATAAACCTGCTAAAGCGAATGTTTTCTTTGCTAATATCCAGTG
K1_K2O5-UTRm_2R	GTTTTCTCTAAAGAAGCAAATACTGGTAAGAATAAACCTGCTAAAGCGAATGTTTTCTTTGCTAATATCCAGTCGTGTACCA ATCATTATAGGAAGTTC
K1_K2O5-UTRres_2R	GAAGCAAATACTGGTAAGAATAAACCTGCTAAAGCGAATGTTTTCTTTGCTAATATCCAGTCGTGTACCAATCATTATAGGA ACACGTCTGTCAATTGTATATAGCAC
K1_ORF4-HA_F2	TATATGATCGTATTAGATATACCTAAAATGAAAAAGCCTGAACTACCGCGAC
K1_ORF4-HA_R1	TTTAGGTATATCTAATACGATCATATATCAAGCATAATCTGGAACATCATATGGATATACACATTTTTCATTCTGTAGA
K1_ORF4-HA_R2	GTCTATTTTACACTTTTGACCTATAAGTCATTTTATTATACACATTTTTCATTATTCTTTGCCCTCGGACGAGTG
K1_UCRmut-O4_R1	ATATGATATATCTTCATTTTAAATAAAAATGTGTTTGAATTAGATAATAATC
K1_UCRmut-O4_R2	AAACACTAAATATATGATATATCTTCATTTTAAATAAAAATGTGGGTAGAATTAGATAATAATCTTTCAGATTTATTCCTTTGC CCTCGGACGAGTG
K1_VV_A27L-term	AAGCAAATACTGGTAAGAATAAACCTGCTAAAGCGAATGATTAAACAAAGAGTTAAGTTACTCATATGGGCGCCGTCCAGT CTGAACATCAATCTTTTTAGAAAACTCATCGAGCATC
K10	AAGTCCAGATAATACTATAG
K1O2_prom_del_F	ATAAGACCGTGAAAGCTACTTAATATATGAAAGTTTTATAATAATTATAACATGGGTAAGGAAAAGACTCACG
K1O2_prom_del2_F	ATAAGACCGTGAAAGCTACTTAATATATGAAAGTTTTATAATAATTATACCATGGGTAAGGAAAAGACTCACG
K2_ORF3_for_seq	ATGAAGAATTTTCAGAAAGAAGTCC
K27	AATGGCTAATAAACAGGCAG
k2orf6primex	CAACATTTCTAAAATAATCGTTAT
KanR1	ATGGGTAAGGAAAAGACTCACG
kanR2	TTAGAAAACTCATCGAGCATCA
KanV2F	GTTGTATTGATGTTGGACGAGTCGG
KanVerF	AGACCGATACCAGGATCTTGCCATC
Kl_Cox1-for2	AAAACCACCAATTAATGCCGG
Kl_Cox1-rev2	GATTAGAGTTAGCCGACCAG
KL_orf6_split1R	AAAACTTTTCATATATTAAGTAGCTTTACGGTCTTAATCATTTTTTATTAATATATCTC
KL_orf6_split2R	GAATTATCGAATTCTACTATGGAATACTTCATTTTAAATTAATAATGTAATTCATATATTATAGTTTAGAAAACTCATCGA GCATC
KL_orf6C_Flag2F	GACCGTGAAAGCTACTTAATATATGAAAGTTTTT
KL_orf6N_EGFP1R	CTTCACCTTTAGACATTTTAGGTATATCTAATACGATCATATAGCTAATGATATTAGAAAACTCATCGAGCATCAA
KL_orf6N_EGFP2F	TATATGATCGTATTAGATATACCTAAAATGTCTAAAGGTGAAGAATTATTCAGTGGTG
KL_orf6N_EGFP2R	CAACATTTCTAAAATAATCGTTATAAAATTTCTATTTGTCCATAATCTGCAGCCCGGGGGATCCACTAGTT
KL_orf6N_Flag1F	AGAAGTTGAATTTGACAACTATCATTAGCTATATGATCGTATTAGGACCGTGAAAGCTACTTAATATATGAAAG
KL_orf6N_Flag1R	CATCTTTATAATCCATTTTAGGTATATCTAATACGATCATATAGCTAATGATATTAGAAAACTCATCGAGCATCAA
KL_orf6N_Flag2F	TATATGATCGTATTAGATATACCTAAAATGGATTATAAAGATGATGATGATAAAATTGATTATGGACAAATAGAAATTTA
KL_orf7C_cMyc1F	TTAGAAAAATATGTAATGAAATGTACTTATAATTATAAATCTTCTTCACTTATTAATTTTTGTTCTAAAAAGAATTTTTCTC GCTAGGGT
KL_orf7C_Flag1F	TTAGAAAAATATGTAATGAAATGTACTTATAATTATTTATCATCATCTTTATAATCATTTAAAAAGAATTTTTCTCGCT AGGGT
KL_orf7C_Flag1R	AGCACTCGTCCGAGGGCAAAGGAATAATGTGAAGTTGATGATATAAAGTAAAAATGAATGAAATATTATTTCTAA
KL_orf7C_Flag2F	TTATTCCTTTGCCCTCGGACGAGTG

KL_orf7C_Flag2R	TTATGTGAAGTTGATGATATAAGTAACTATAATATATGAATTACATTATTAATTTAAAAATGAAAAAGCCTGAACTCACC GCGAC
olig2(dC)anchor	GACCACGCGTATCGATGTCGACCCCCCCCCC
oligo(dG)anch2	GATTGAGGTGATCTGATGTCGAGGGGGGGGGGG
oligod(T)anchor	GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTT
ORF11_K2_tail_2	TGCCTCTACAAAATAATAAAATAATTCTTG
ORF1F2r	TTATTCACACATCTAGGTTGAGGGA
ORF1-K1_tail_2	TGGGACACTGATTTTATGCTG
ORF1-K1_tail_3	AGGATCAGAAGTAGGACAATTAGAAT
ORF1-K2_tail_2	CCTCCATTCTATTTATGGCATAAAAG
ORF2-pGKL1_tail	TGAGGTCTTAGAAAGTGTTGGTTG
ORF3_k1_tail_2	TTGTGTCTCTTATAGGCCCTT
ORF3_SAMdel_F1	ATTAAGTATGCTTCCTAAAACCACTTTATTGGCTATTGGGTCTGCAAAAGCTGGAGCTATACTAAATGGACTAATCTAA
ORF3_SAMdel_R1	AAAAACTTTTCATATATTAAGTAGCTTTCACGGTCTCATTTTTTAGAAAAGAAATGA
ORF3_SAMdel_R2	TTAAGAATGCTAATTCATCATTCTTTTTAAATTAATAATGTAATTCATATATTATAGTTTAGAAAACTCATCGAGCATCAAA
ORF3-K2_tail_3	AAGAAGATTGCTTCGAATGTAAAATTG
ORF3-pGKL1_tail	ACTTATCGATTTTCCACGCC
ORF3-pGKL2_tail	TGAAAAAGTTTATGCAGTAGATCCA
ORF4_K2_tail_2	TCTACATGCAGCACACTATGA
ORF4-EGFP_1F	CAAAATAGATATTTTGACAAACAAATTGAATATGATTTAGGAGGAGGGAATGTCTAAAGGTGAAGAATTATTC
ORF4-EGFP_1R	AAAAACTTTTCATATATTAAGTAGCTTTCACGGTCTTATTGTACAATTCATCCATACC
ORF4-EGFP_2R	CTAATTTAATTAATTTGGTGCAATTTTGCATAAATCATTTTTAAATTAATAATGTAATTCATATATTATAGT
ORF4-K2_tail_3	CGAATAAATTCAAAATTTCTTGGTGC
ORF4-pGKL1_tail	AGAAGCTCTAATAGTGTTGAAGG
ORF6+7join_F1	CAGGTTATTTAGAAAATATATGTAATGAAATGTACTTAGATATGGTGGTCGACGGATAC
ORF6+7join_F2	CAATTATACTCTAGGTGGTGGAACATAATGAAAATATTATTTCTAATTTAA
ORF6+7join_R1	TTAAATTAGAAATAATATTTTCATTAGTTCACCACCTAGAGTATAATTG
ORF6+7join_R2	AAAAACTTTTCATATATTAAGTAGCTTTCACGGTCTTATAAAAAGAAATATTTCTCGCTAG
ORF6+7join_R3	CAATTTTATGTGAAGTTGATGATATAAAGTAATTAGAAAACTCATCGAGCATCAAA
ORF6-K2_tail_2	CGTTGTTGTGCAATCATTTAATGCAG
ORF7-K2_tail_2	GATTCCACTTTAGAAGACGAAGTAGAAG
ORF7-pGKL2_tail	TTTCTCAGAGATAACCCAGAGAAC
ORF8-pGKL2_tail	CCAATGTTTCATTGTGCGAATACT
ORF9_D13_F1	ATCTTATACATTATAGAGATTATTGAGCATAAAGTAGAATTCAAATATGCCAGAAGTTAAAGGTATGGGAAG
ORF9_D13_R1	AAAAACTTTTCATATATTAAGTAGCTTTCACGGTCTTAGTTATTATCTCCATAATCTTG
ORF9_D13_R2	TGAGTTTTCTGCCTGTTTATTAGCCATTTTTAAATTAATAATGTAATTCATATATTATAGTTTAGAAAACTCATCGAGCAT CAAA
ORF9-K2_tail_2	TGTAGTAAACAAAAAGATATGGCAG
ORF9Rsekven	CCAGCTAAGTTTTTCCATTCTG
PCR-anchor	GACCACGCGTATCGATGTCGAC
pGEX_dual_F	CCAGGATCCGTCGACGTTTAACTTTAAGAAGGAGATATACCATGAATGAAAATATTATTTCT
pGEX_dual_His_R	CCAGGATCCGCGGCCGCTTAGTGATGGTGATGGTGATGGTGATGGTGACCTAAAAAGAAATATTTCTCGCTAG
pGEX_dual_R	CCAGGATCCGCGGCCGCTTATAAAAAGAAATATTTCTCGCTAG
pGKL_D12L_prom_F	GAATAATTTAGAAGTTGAATTTGACAACTATCATTAGCTATTTAATTAGAAAACTCATCGAGCATC
pGKL_D12L_prom_R	GTCCATAATCCATTTTAGGTATATCTAATACGATCATATAAATAATGAAAACAACTATAGAGTTGTAAATGGGTAAGGAAA AGACTCAC
pGKL_HA-ORF4_F1	AACAATCAGATTAATAAAGTAAAGTATCACTAAATATATGATATATCTTCATTTAATTAATAAATGAAAAAGCCTGAACTCA CCGCGAC
pGKL_HA-ORF4_F2	CTATCTGATTATAACAATCAGATTAATAAATGTATCCATATGATGTTCCAGATTATGCTGGTAAAGTATCACTAATAAAAAAT
pGKL_HA-ORF4_R1	TTTTAATCTGATTGTTATAATCAGATAGTTATTCCTTTGCCCTCGGACGAGTG
pGKL_HA-ORF6_F1	GAAGTTGAATTTGACAACTATCATTAGCTATATGAATATATGATATATCTTCATTTTAAATTAATAAATGAAAAAGCCTGAACT CACCAGCGAC
pGKL_HA-ORF6_F2	TATATGATCGTATTAGATATACCTAAAATGTATCCATATGATGTTCCAGATTATGCTGATTATGGACAAATAGAAATTTA
pGKL_HA-ORF6_R1	TTTAGGTATATCTAATACGATCATATATTCTTCTTCCCTCGGACGAGTG
pGKL_HA-ORF9_F1	GAGATTATTGAGCATAAAGTAGAATTCAAATATAAAAATATATGATATATCTTCATTTTAAATTAATAAATGAAAAAGCCTGAA CTACCCGCGAC
pGKL_HA-ORF9_F2	CTATCTGATTATAACAATCAGATTAATAAATGTATCCATATGATGTTCCAGATTATGCTGGTGAGATTATTGAGCATAAAGTA G

pGKL_HA-ORF9z_F1	AGTATTCGACAAATGAAACATTGGTTATCTTATACATTTTTAAATTAATAATGAATTCATATATTATAGTTTATTCCTTTGCCCTCGGACGAGTG
pGKL_HA-ORF9z_R1	TTTTAATCTGATTGTTATAATCAGATAGATATATGATATATCTTCATTTTAATTAATAATGAAAAAGCCTGAACTCACCGGACGAC
pGKL_His-ORF6_F2	TATATGATCGTATTAGATATACCTAAAAATGCATCATCACCATCACCATGATTATGGACAAATAGAAATTTA
pGKL_I4L_prom_R	GTCCTAATCCATTTTAGGTATATCTAATACGATCATATAATTAATGAAAAGTTAAATAATTTTTTATTAAATGGGTAAAGGAAAGAACTCAC
pGKL_ORF3-HA_F1	TTAAGAATGCTAATTCATCATTCTTTTTAAATTAATAATGAATTCATATATTATAGTTTATTCCTTTGCCCTCGGACGAGTG
pGKL_ORF3-HA_F2	TTTTAATTAATAATGAAGATATATCATATATTCAGCATAATCTGGAACATCATATGGATATCCTTTTTAGAAAAAGAAATGATAAG
pGKL_ORF3-HA_R1	ATATATGATATATCTTCATTTTAATTAATAATGAAAAAGCCTGAACTCACCGGACGAC
pGKL_ORF3-NLS_F2	TTTTAATTAATAATGAAGATATATCATATATTCAAACTTTTCTTTTTTTTTTGGTCTTTTTAGAAAAAGAAATGATAAG
pGKL_ORF9-HA_R1	TTTTAATTAATAATGAAGATATATCATATATTCAGCATAATCTGGAACATCATATGGATATCCTTCTGATATTTGAAAGTTATTAATC
pGKL_ORF9-HA_R2	GTTTTCTGCCTGTTTATTAGCCATTTTAAATTAATAATGAATTCATATATTATAGTTTATTCCTTTGCCCTCGGACGAGTG
pGKL_VV_prom3_R	TGATGCTCGATGAGTTTTCTAATTATTCCTTTGCCCTCGGACGAGTG
pGKL1_HA-ORF1_F1	TTTACACTTTTGACCTATAAGTCATTTTATTATACACATTTTCCAATATATGATATATCTTCATTTTAATTAATAATGAAAAAGCCTGAACTCAC
pGKL1_HA-ORF1_F2	ACTATAATATATGAATTACATTATTAATTTAAAAATGTATCCATATGATGTTCCAGATTATGCTGATTACAAAGATAAGGCTTTAAAT
pGKL1_HA-ORF1_R1	TTTTAAATTAATAATGAATTCATATATTATAGTTTATTCCTTTGCCCTCGGACGAGTG
pGKL2_GST-ORF6_F2	TATATGATCGTATTAGATATACCTAAAAATGTCCCCTATACTAGGTTATTG
pGKL2_GST-ORF6_R1	CATTTTAGGTATATCTAATACGATCATATATTAGAAAAACTCATCGAGCATCAA
pGKL2_ORF3-EGFP_F1	CATCTATTTGAGAATTTAAGAATGCTAATTCATCATTCTTTTTAAATTAATAATGAATTCATATATTATAGTTTAGAAAAACTCATCGAGCATC
pGKL2_ORF3-EGFP_F2	TTCATATATTAAGTAGCTTTACGGTCTCACTGCAGCCCCGGGGATCCAC
pGKL2_ORF3-EGFP_R2	CAAGAAATAGAATTATCACGCATGTATTCTTATCATTTCTTTCTAAAAAATCTAAAGGTGAAGAATTATTC
pGKL-kanF	ATAAGACCGTGAAAGCTACTTAATATATGAAAGTTTTATAATAATTATAAAATGGGTAAAGGAAAGACTCACG
pGKL-kanR	GTTTTCTCTAAAGAAGCAATACTGTAAGAATAAACCTGCTAAAGCGAATGTTAGAAAAACTCATCGAGCATCA
Sc/Kl_Cox1-for	ATCAATGGGTGCTATTTCTC
Sc/Kl_Cox1-rev	GCTGGTGGAGAAGTTAATAAG
Sc/Kl_Cox2-for	AAGGTATTTAGAATTACATG
Sc/Kl_Cox2-rev	TCAGGAATAACATATGATTC
T3	AATTAACCTCACTAAAGGG
T7	GTAATACGACTCACTATAGGGC
vORF3-k2-forw	AAATCTGGAATCTGCTATGG
vORF3-k2-rev_2	GTTCTTTTGTAGCCGGTATT

## 4.1.9 Protílátky

Název	Typ	Výrobce
anti-c-Myc (9E10)	myší monoklonální IgG <sub>1</sub>	Santa Cruz Biotechnology, Inc.
anti-eIF4G1	králičí sérum	získáno od Leoše S. Valáška (MBÚ)
anti-FLAG <sup>®</sup> (M2)	myší monoklonální IgG <sub>1</sub>	Sigma
anti-GFP (B-2)	myší monoklonální IgG <sub>2a</sub>	Santa Cruz Biotechnology, Inc.
anti-GST (3D4)	myší monoklonální IgG <sub>2b</sub>	Santa Cruz Biotechnology, Inc.
anti-HA (6E2)	myší monoklonální IgG <sub>1</sub>	Cell Signaling
anti-HA (F-7)	myší monoklonální IgG <sub>2a</sub>	Santa Cruz Biotechnology, Inc.
anti-HA (HA-7)	myší monoklonální IgG <sub>1</sub>	Sigma
anti-mouse IgG-HRP	kozí polyklonální	Santa Cruz Biotechnology, Inc.

anti-polyHis (HIS1)	myší monoklonální IgG <sub>2a</sub>	Sigma
anti-rabbit IgG-HRP	kozí polyklonální	BIO-RAD

#### 4.1.10 Bakteriální kmeny

Tabulka 8: Bakteriální kmeny použité v této práci		
Název kmene	Genotyp	Zdroj / Původ
<i>Escherichia coli</i>		
Rosetta <sup>TM</sup> (DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub>(R<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pRARE (Cam<sup>R</sup>)</i>	Novagen
Rosetta 2(DE3)pLysS	F <sup>-</sup> <i>ompT hsdS<sub>B</sub>(R<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pLysSpRARE2 (Cam<sup>R</sup>)</i>	Novagen
XL-1 Blue	<i>endA1 gyrA96(nal<sup>R</sup>) thi-1 recA1 relA1 supE44 lac [F' Tn10(tet<sup>R</sup>) proAB lacI<sup>Δ</sup>(lacZ)M15] hsdR17(r<sub>K</sub><sup>-</sup> m<sub>K</sub><sup>+</sup>)</i>	Stratagene

#### 4.1.11 Kvasinkové kmeny

Tabulka 9: Kvasinkové kmeny použité v této práci		
Název kmene	Genotyp	Zdroj / Původ
<i>Kluyveromyces lactis</i>		
IFO1267	MATa pGKL1 <sup>+</sup> pGKL2 <sup>+</sup>	Institut pro fermentaci, Osaka
IFO1267_pGKL <sup>-</sup>	MATa	Tato práce
IFO1267_pRKL1-1	MATa pRKL1-1 <sup>+</sup> (K1UCR2-G418 <sup>R</sup> , orf2Δ) pGKL2 <sup>+</sup>	V. Vopálenský
IFO1267_pRKL1-2	MATa pRKL1-2 <sup>+</sup> (K1UCR2*-G418 <sup>R</sup> , orf2Δ) pGKL2 <sup>+</sup>	Tato práce
IFO1267_pRKL1-3	MATa pRKL1-3 <sup>+</sup> (K1UCR2**-G418 <sup>R</sup> , orf2Δ) pGKL2 <sup>+</sup>	Tato práce
IFO1267_pRKL1-4/2-4	MATa pRKL1-4 <sup>+</sup> (ORF4-HA, HygB <sup>R</sup> ) pRKL2-4 <sup>+</sup> (yEGFP3-ORF6, G418 <sup>R</sup> )	Tato práce
IFO1267_pRKL1-5	MATa pRKL1-5 <sup>+</sup> (G418 <sup>R</sup> -K2ORF5 3' UTR, orf2Δ) pGKL2 <sup>+</sup>	Tato práce
IFO1267_pRKL1-6	MATa pRKL1-6 <sup>+</sup> (G418 <sup>R</sup> -K2ORF5 3' UTR°, orf2Δ) pGKL2 <sup>+</sup>	Tato práce
IFO1267_pRKL1-7	MATa pRKL1-7 <sup>+</sup> (G418 <sup>R</sup> -K2ORF5 3' UTR°, orf2Δ) pGKL2 <sup>+</sup>	Tato práce
IFO1267_pRKL1-8	MATa pRKL1-8 <sup>+</sup> (G418 <sup>R</sup> -A27L 3' UTR, orf2Δ) pGKL2 <sup>+</sup>	Tato práce
IFO1267_pRKL1-9	MATa pRKL1-9 <sup>+</sup> (K1UCR4**-ORF4, HygB <sup>R</sup> ) pGKL2 <sup>+</sup>	Tato práce
IFO1267_pRKL1-10/2-4	MATa pRKL1-10 <sup>+</sup> (HA-ORF3, HygB <sup>R</sup> ) pGKL1 <sup>+</sup> pRKL2-4 <sup>+</sup> (yEGFP3-ORF6, G418 <sup>R</sup> )	Tato práce
IFO1267_pRKL2-1	MATa pGKL1 <sup>+</sup> pRKL2-1 <sup>+</sup> (G418 <sup>R</sup> )	Tato práce
IFO1267_pRKL2-2	MATa pGKL1 <sup>+</sup> pRKL2-2 <sup>+</sup> (ORF3ΔSAM, G418 <sup>R</sup> )	Tato práce
IFO1267_pRKL2-3	MATa pGKL1 <sup>+</sup> pRKL2-3 <sup>+</sup> (Flag-ORF6, G418 <sup>R</sup> )	SÝKORA (2013)
IFO1267_pRKL2-4	MATa pGKL1 <sup>+</sup> pRKL2-4 <sup>+</sup> (yEGFP3-ORF6, G418 <sup>R</sup> )	SÝKORA (2013)
IFO1267_pRKL2-5	MATa pGKL1 <sup>+</sup> pRKL2-5 <sup>+</sup> (yEGFP3-ORF6, ORF7-Flag, G418 <sup>R</sup> , HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-6	MATa pGKL1 <sup>+</sup> pRKL2-6 <sup>+</sup> (yEGFP3-ORF6, ORF3-HA, G418 <sup>R</sup> , HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-7	MATa pGKL1 <sup>+</sup> pRKL2-7 <sup>+</sup> (yEGFP3-ORF6, HA-ORF4, G418 <sup>R</sup> , HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-8	MATa pGKL1 <sup>+</sup> pRKL2-8 <sup>+</sup> (yEGFP3-ORF4, G418 <sup>R</sup> )	Tato práce
IFO1267_pRKL2-9	MATa pGKL1 <sup>+</sup> pRKL2-9 <sup>+</sup> (yEGFP3-ORF4, HA-ORF6, G418 <sup>R</sup> , HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-10	MATa pGKL1 <sup>+</sup> pRKL2-10 <sup>+</sup> (yEGFP3-ORF4, ORF3-HA, G418 <sup>R</sup> , HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-11	MATa pRKL2-11 <sup>+</sup> (ORF3-yEGFP3, G418 <sup>R</sup> )	Tato práce
IFO1267_pRKL2-12	MATa pGKL1 <sup>+</sup> pRKL2-12 <sup>+</sup> (HA-ORF6, HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-13	MATa pGKL1 <sup>+</sup> pRKL2-13 <sup>+</sup> (HA-ORF4, HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-14	MATa pGKL1 <sup>+</sup> pRKL2-14 <sup>+</sup> (ORF3-HA, HygB <sup>R</sup> )	Tato práce

IFO1267_pRKL2-15	MATa pGKL1 <sup>+</sup> pRKL2-15 <sup>+</sup> (ORF7-Flag, HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-16	MATa pGKL1 <sup>+</sup> pRKL2-16 <sup>+</sup> (yEGFP3-ORF6, ORF7-c-Myc, G418 <sup>R</sup> , HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-17	MATa pGKL1 <sup>+</sup> pRKL2-17 <sup>+</sup> (orf9::D13, G418 <sup>R</sup> ) pGKL2 <sup>+</sup>	Tato práce
IFO1267_pRKL2-18Δ	MATa pGKL1 <sup>+</sup> pRKL2-18Δ <sup>+</sup> (G418 <sup>R</sup> ) pGKL2 <sup>+</sup>	Tato práce
IFO1267_pRKL2-19	MATa pGKL1 <sup>+</sup> pRKL2-19 <sup>+</sup> (ORF6β, ORF6β', orf6Δ, G418 <sup>R</sup> ) pGKL2 <sup>+</sup>	Tato práce
IFO1267_pRKL2-20	MATa pRKL2-20 <sup>+</sup> (HA-ORF9, HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-21	MATa pRKL2-21 <sup>+</sup> (HA-ORF9, HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-22	MATa pGKL1 <sup>+</sup> pRKL2-22 <sup>+</sup> (yEGFP3-ORF6, ORF9-HA, HygB <sup>R</sup> , G418 <sup>R</sup> ) pRKL2-4 <sup>+</sup>	Tato práce
IFO1267_pRKL2-23	MATa pGKL1 <sup>+</sup> pRKL2-23 <sup>+</sup> (yEGFP3-ORF6, HA-ORF9, HygB <sup>R</sup> , G418 <sup>R</sup> ) pRKL2-4 <sup>+</sup>	Tato práce
IFO1267_pRKL2-24	MATa pRKL2-24 <sup>+</sup> (ORF3-yEGFP3, HA-ORF9, HygB <sup>R</sup> , G418 <sup>R</sup> )	Tato práce
IFO1267_pRKL2-25	MATa pGKL1 <sup>+</sup> pRKL2-25 <sup>+</sup> (His-ORF6, HygB <sup>R</sup> )	Tato práce
IFO1267_pRKL2-26	MATa pGKL1 <sup>+</sup> pRKL2-26 <sup>+</sup> (GST-TEV-ORF6, G418 <sup>R</sup> )	Tato práce
IFO1267_pRKL2-27	MATa pGKL1 <sup>+</sup> pRKL2-27 <sup>+</sup> (HygB <sup>R</sup> , D12Lprom-G418 <sup>R</sup> )	Tato práce
IFO1267_pRKL2-28	MATa pGKL1 <sup>+</sup> pRKL2-28 <sup>+</sup> (ORF4-yEGFP3, G418 <sup>R</sup> ) pGKL2 <sup>+</sup>	Tato práce
<i>Saccharomyces cerevisiae</i>		
S6/1	MATα	WOODS a BEVAN (1968)
F102-2	MATα leu2-3,112 his4-519 met15 can1 p <sup>0</sup> dsRNA-L pGKL1 <sup>+</sup> pGKL2 <sup>+</sup>	GUNGE <i>et al.</i> (1981)
F102-2_pRKL2-21	MATα leu2-3,112 his4-519 met15 can1 p <sup>0</sup> dsRNA-L pGKL1 <sup>+</sup> pRKL2-21 <sup>+</sup> (HA-ORF9, HygB <sup>R</sup> ) pGKL2 <sup>+</sup>	Tato práce

UCR - sekvence mezi iniciačním kodónem a UCS (včetně) příslušného ORF

\* - UCR sekvence s jednou bodovou mutací v předpokládaném iniciátorovém regionu (INR)

\*\* - UCR sekvence se dvěma bodovými mutacemi v předpokládaném INR

° - 3' nepřekládaná oblast genu *K2ORF5* s mutacemi v předpokládané RNA vlásence 2

°° - 3' nepřekládaná oblast genu *K2ORF5* s mutacemi obnovujícími předpokládanou RNA vlásenku 2

Detailní popis všech plasmidů je uveden v kapitole 4.1.7.

## 4.1.12 Kultivační média

### 4.1.12.1 Bakteriální kultivační média

#### 4.1.12.1.1 Tekutá média

**2x TY médium:** 1,6 % (w/v) pepton; 1 % (w/v) kvasničný autolyzát; 0,5 % (w/v) NaCl;  
upravit pomocí NaOH na pH 7,0

**2x TY médium s antibiotiky:** složení stejné jako 2x TY médium; přidána antibiotika v celkové koncentraci: ampicilin (100 µg/ml), kanamycin (50 µg/ml), chloramfenikol (34 µg/ml)

#### 4.1.12.1.2 Agarové půdy

**2x TYA:** 2x TY médium + 2 % (w/v) agar

**2x TYA s antibiotiky:** 2x TY médium s antibiotiky + 2 % (w/v) agar

## 4.1.12.2 Kvasinková kultivační média

### 4.1.12.2.1 Tekutá média

**SD–:** 0,174 % (w/v) *Yeast Nitrogen Base*; 0,5 % (w/v)  $(\text{NH}_4)_2\text{SO}_4$ ; 0,2 % (w/v) glukosa

**SD+:** SD–; 0,003 % (w/v) adenin (A); 0,003 % (w/v) L-histidin (H);

0,003 % (w/v) L-leucin (L); 0,003 % (w/v) L-lysin (Lys);

0,003 % (w/v) L-methionin (M); 0,003 % (w/v) L-tryptofan (T);

0,003 % (w/v) uracil (U)

**selekční minimální média:** složení jako SD+, ale chybí zde konkrétní aminokyselina, báze nebo kombinace více chybějících složek

**YPD:** 0,5 % (w/v) kvasniční autolyzát; 1 % (w/v) pepton; 2 % (w/v) glukosa

**YPD + G418:** YPD; G418 (250 µg/ml)

**YPD + HygB:** YPD; HygB (200 µg/ml)

**YPD + G418 + HygB:** YPD; G418 (250 µg/ml); HygB (200 µg/ml)

### 4.1.12.2.2 Agarové půdy

**SD půdy:** SD–; SD+ nebo selekční média + 2 % (w/v) agar

**SD– + G418:** SD– + 2 % (w/v) agar; G418 (100–500 µg/ml)

**YPDA:** YPD + 2 % (w/v) agar

**YPDA + G418:** YPDA; G418 (250 µg/ml nebo 750 µg/ml)

**YPDA + HygB:** YPDA; HygB (200 µg/ml nebo 300 µg/ml)



## **4.2 METODY**

### **4.2.1 Práce s mikroorganismy**

Veškerá práce s mikroorganismy probíhala v UV boxu.

#### **4.2.1.1 Kultivace bakterií**

Kultivace bakteriálních kmenů probíhala v tekutém médiu za aerobních podmínek a třepání v reciproké nebo orbitální třepačce v Erlenmayerových baňkách, nebo na agarových půdách v termostatu. Pokud není uvedeno jinak, kultivace probíhala při teplotě 37 °C.

#### **4.2.1.2 Kultivace kvasinek**

Kultivace kvasinkových kmenů probíhala v tekutém médiu za aerobních podmínek a třepání v reciproké nebo orbitální třepačce v Erlenmayerových baňkách, nebo na agarových půdách v termostatu. Kultivace probíhala při teplotě 28 °C.

#### **4.2.1.3 Skladování bakteriálních a kvasinkových kmenů**

Krátkodobé skladování bakteriálních kmenů v řádu týdnů a kvasinkových kmenů v řádu týdnů až měsíců probíhalo na agarových půdách při teplotě 4 °C. Pro dlouhodobé skladování byly čerstvě napěstované mikroorganismy na agarových půdách převedeny do 50 % (v/v) glycerolu a skladovány při –80 °C.

#### **4.2.1.4 Měření růstu mikroorganismů**

Měření průběhu růstu bakterií a kvasinek probíhalo měřením změny optické denzity (OD) na přístroji Spectronic® Helios Beta (Thermo Spectronic). Pro měření bylo použito záření o vlnové délce 600 nm nebo 660 nm, tato hodnota je dále zapisována jako dolní index. Pokud hodnota OD přesahovala hodnotu 1, byl vzorek před měřením naředěn.

##### **4.2.1.4.1 Stanovení růstové křivky kvasinkového kmene**

Kultura kvasinkového kmene narostlá přes noc v tekutém médiu při 28 °C byla zaočkována do vytemperovaného média na OD<sub>600</sub> 0,05 nebo 0,1. V pravidelných intervalech byly odebírány vzorky a měřeny hodnoty optické denzity při 600 nm. Tyto hodnoty byly vyneseny do grafu v semilogaritmickém zobrazení.

#### 4.2.1.5 Indukce bakteriálních promotorů s *lac* operátorem

Vektor plasmidového typu, kódující příslušný konstrukt čerstvě vnést do elektrokompetentních buněk *E. coli* Rosetta<sup>TM</sup>(DE3) nebo Rosetta 2(DE3)pLysS elektroporací, buňky vysít na selekční médium obsahující chloramfenikol a ampicilin.

- pomocí mikrobiologické kličky část biomasy přenést do selekčního média a kultivovat v orbitální třepačce přes noc při 28 °C
- kulturu přeočkovat na OD<sub>600</sub> 0,05–0,1 do požadovaného množství selekčního média vytemperovaného na 28 °C a kultivovat do dosažení OD<sub>600</sub> 0,35–0,4
- snížit kultivační teplotu na 15 °C po dobu 3–5 hod.
- přidat IPTG na finální koncentraci 0,1 mM a inkubovat přes noc při 15 °C
- centrifugovat 10 min./5000g při 4 °C, odstranit supernatant
- buňky propláchnout v ledovém TKG pufru, ihned použít nebo uchovat při –80 °C

### 4.2.2 Práce s DNA

#### 4.2.2.1 Izolace DNA

##### 4.2.2.1.1 Minipreparace vektorů plasmidového typu z buněk *E. coli*

- bakteriální kulturu nechat narůst přes noc v termostatu při teplotě 37 °C na agarových půdách 2x TYA obsahujících příslušné selekční antibiotikum
- sterilním párátkem přenést množství bakteriální biomasy odpovídající ploše 2 x 2 cm do 400 µl roztoku STET, přidat 5 µl 5 % (w/v) lysozymu a dokonale resuspendovat
- inkubovat minimálně 2 min. v termobloku při 95 °C
- nechat vychladnout 5–10 min. při pokojové teplotě
- centrifugovat 10 min./14000g při pokojové teplotě
- odstranit pelet párátkem
- přidat 400 µl isopropanolu
- inkubovat 30–60 min. při teplotě –20 °C
- centrifugovat 15 min./14000g při pokojové teplotě
- odstranit supernatant, sediment resuspendovat v 1 ml 70 % (v/v) ethanolu
- centrifugovat 10 min./14000g při pokojové teplotě
- odstranit supernatant, pelet vysušit 60 min. v termostatu při teplotě 37 °C nebo 10–15 min. v přístroji Speed Vac Plus SC110A (Savant)
- vysušený pelet resuspendovat v 30 µl roztoku TE
- 5 µl vzorku analyzovat elektroforeticky

#### 4.2.2.1.2 Minipreparace vektorů plasmidového typu z buněk *E. coli* pomocí soupravy

##### **GenBond Plasmid FlexSpin Kit či FavorPrep™ Plasmid Extraction Mini Kit**

- bakteriální kulturu nesoucí příslušný vektor plasmidového typu nechat narůst přes noc při teplotě 37 °C v 2–5 ml média 2xTY obsahujícího příslušné selekční antibiotikum nebo v ploše 2 x 3 cm na agarových půdách 2x TYA obsahujících příslušné selekční antibiotikum
- centrifugovat 1 min./14000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat pipetováním v 200 µl pufru S1 či 200 µl pufru FAPD1 (obsahují ribonukleázu A)
- přidat 200 µl pufru S2 či 200 µl pufru FAPD2, opatrně promíchat otáčením mikrozkmavky, inkubovat 2–4 min. při pokojové teplotě
- přidat 350 µl pufru S3 či 300 µl pufru FAPD3, opatrně promíchat otáčením mikrozkmavky, inkubovat 1 min. při pokojové teplotě
- centrifugovat 5 min./14000g při pokojové teplotě
- supernatant přenést do dodaných kolonek, přidat 100 µl DNA vazebné matrice (v případě soupravy GenBond Plasmid FlexSpin Kit)
- centrifugovat 1 min./14000g při pokojové teplotě
- odstranit proteklý supernatant, přidat 700 µl pufru WB či 400 µl pufru W1
- centrifugovat 1 min./14000g při pokojové teplotě
- odstranit proteklý supernatant, přidat 250 µl pufru WB či 700 µl pufru Wash Buffer
- centrifugovat 1 min./14000g při pokojové teplotě
- odstranit proteklý supernatant, centrifugovat 2 min./18000g při pokojové teplotě
- kolonku nasadit na čistou mikrozkmavku
- přidat 50–100 µl ddH<sub>2</sub>O zahřáté v termobloku na 60 °C, inkubovat 1 min. při pokojové teplotě
- centrifugovat 2 min./14000g při pokojové teplotě
- 5 µl vzorku analyzovat elektroforeticky

#### 4.2.2.1.3 Midipreparace vektorů plasmidového typu z buněk *E. coli*

- bakteriální kulturu nesoucí příslušný vektor plasmidového typu nechat narůst přes noc při teplotě 37 °C v 50 ml média 2xTY obsahujícího příslušné selekční antibiotikum
- centrifugovat 10 min./5000g při pokojové teplotě
- odstranit supernatant, pelet důkladně resuspendovat v 2 ml roztoku I
- přidat 4 ml roztoku II, promíchat otáčením zkumavky, inkubovat 5 min. při pokojové teplotě
- přidat 3 ml roztoku III, promíchat otáčením zkumavky, inkubovat 20 min. na ledu
- centrifugovat 20 min./5000g při teplotě 4 °C
- supernatant přefiltrovat přes gázu, přidat 0,7 objemu isopropanolu, dokonale promíchat, inkubovat 10 min. na ledu
- centrifugovat 20 min./5000g při teplotě 4 °C

- odstranit supernatant, pelet resuspendovat v 3 ml 70 % (v/v) ethanolu
- centrifugovat 5 min./5000g při pokojové teplotě
- dokonale odstranit supernatant, pelet nechat vyschnout v termostatu při teplotě 37 °C
- pelet resuspendovat v 500 µl pufru TE, přidat 10 µl roztoku TRN, inkubovat 1 hod. v termostatu při teplotě 37 °C
- přidat 500 µl upraveného fenolu pH 8, dokonale promíchat
- centrifugovat 5 min./14000g při pokojové teplotě
- odebrat vrchní vodnou fázi, doplnit její objem do 500 µl TE pufrém a přidat k ní 250 µl upraveného fenolu a 250 µl upraveného chloroformu, dokonale promíchat
- centrifugovat 5 min./14000g při pokojové teplotě
- opakovat předchozí dva kroky 1x
- odebrat vrchní vodnou fázi, doplnit její objem do 500 µl TE pufrém a přidat k ní 500 µl upraveného chloroformu, dokonale promíchat
- centrifugovat 5 min./14000g při pokojové teplotě
- odebrat vrchní vodnou fázi, přidat k ní 0,1 objemu 3 M NaAc a 2,5 objemu 96 % (v/v) ethanolu vychlazeného na 4 °C
- inkubovat přes noc při –20 °C
- centrifugovat 10 min./14000g při pokojové teplotě
- odstranit supernatant, přidat 1 ml 70 % (v/v) ethanolu, dokonale promíchat
- centrifugovat 10 min./14000g při pokojové teplotě
- odstranit supernatant, pelet nechat dokonale vyschnout v termostatu při teplotě 37 °C
- podle objemu peletu rozpustit v 100–300 µl pufru TE
- 1 µl vzorku analyzovat elektroforeticky

#### 4.2.2.1.4 Minipreparace kvasinkových cytoplasmatických plasmidů z buněk *K. lactis*

Adaptováno podle (POSPÍŠEK a PALKOVÁ 1991).

- kvasinkovou kulturu nesoucí příslušné plasmidy nechat narůst 1–2 dny v termostatu při teplotě 28 °C na agarových půdách YPDA, YPDA + G418 nebo YPDA + HygB
- množství kvasinkové biomasy objemu špendlíkové hlavičky (~10 µl) přenést párátkem do jamky mikrotitrační destičky
- inkubovat v hybridizační peci při teplotě 45 °C 2–6 hod.
- přidat 40 µl pufru TESP
- inkubovat v třepačce pro mikrotitrační destičky PST-60HL (BIOSAN Ltd.) 15 min./900 otáček za minutu při teplotě 25 °C
- inkubovat přes noc v termostatu při teplotě 37 °C
- přidat 40 µl 1x vzorkového pufru pro agarosovou elektroforézu
- inkubovat v třepačce pro mikrotitrační destičky PST-60HL (BIOSAN Ltd.) 15–30 min./900 otáček za minutu při teplotě 25 °C
- 15 µl vzorku analyzovat elektroforeticky

#### 4.2.2.1.5 Izolace DNA fragmentu z agarosového gelu pomocí soupravy FastBack DNA Minispin Kit či FavorPrep™ GEL/PCR Purification Kit

- vzorky rozdělit na 0,5 % (w/v) agarosovém gelu či 1,5 % (w/v) agarosovém gelu (fragменты pod 500 bp)
- pod transluminátorem generujícím UV záření o vlnové délce 254 nm nebo 312 nm vyříznout oblast gelu s požadovaným fragmentem DNA
- vyříznutou oblast gelu přenést do mikrozkušavky a podle velikosti přidat 500–650 µl dodaného pufru *GL* nebo *FADF*
- inkubovat 5–10 min. v termobloku při teplotě 60 °C za občasného promíchání
- přenést suspenzi do dodané kolonky se skleněnou fritou
- centrifugovat 1 min./14000g při pokojové teplotě
- odstranit proteklý supernatant, přidat 750 µl dodaného pufru *WB* či 700 µl pufru *Wash Buffer*
- centrifugovat 1 min./14000g při pokojové teplotě
- odstranit proteklý supernatant, centrifugovat 3 min./18000g při pokojové teplotě či v případě soupravy FastBack DNA Minispin Kit přidat 250 µl pufru *WB* a centrifugovat 1 min./14000g při pokojové teplotě
- odstranit proteklý supernatant, skleněnou fritu nasadit na čistou mikrozkušavku
- přidat 30–40 µl ddH<sub>2</sub>O zahřáté v termobloku na 60 °C, inkubovat 1 min. při pokojové teplotě
- centrifugovat 2 min./14000g při pokojové teplotě
- 5 µl vzorku analyzovat elektroforeticky

#### 4.2.2.1.6 Izolace produktu PCR reakce pomocí soupravy High Pure PCR Product Purification Kit

- produkt PCR reakce doplnit do objemu 100 µl pomocí ddH<sub>2</sub>O, přidat 500 µl dodaného pufru *binding*, dokonale promíchat
- přenést do dodané kolonky se skleněnou fritou
- centrifugovat 1 min./14000g při pokojové teplotě
- odstranit proteklý supernatant, přidat 500 µl dodaného pufru *wash*
- centrifugovat 1 min./14000g při pokojové teplotě
- odstranit proteklý supernatant, přidat 200 µl dodaného pufru *wash*
- centrifugovat 1 min./14000g při pokojové teplotě
- odstranit proteklý supernatant, skleněnou fritu nasadit na čistou mikrozkušavku
- přidat 50–100 µl ddH<sub>2</sub>O zahřáté v termobloku na 60 °C, inkubovat 1 min. při pokojové teplotě
- centrifugovat 2 min./14000g při pokojové teplotě
- 5–10 µl vzorku analyzovat elektroforeticky

#### 4.2.2.1.7 Izolace komplementární DNA (cDNA) po reverzní transkripci pomocí soupravy High Pure PCR Product Purification Kit

- k 20 µl směsi po reverzní transkripci přidat 100 µl dodaného pufru *binding*, dokonale promíchat
- přenést do dodané kolonky se skleněnou fritou
- centrifugovat 30 s/7000g při pokojové teplotě
- odstranit proteklý supernatant, přidat 500 µl dodaného pufru *wash*
- centrifugovat 30 s/7000g při pokojové teplotě
- odstranit proteklý supernatant, přidat 200 µl dodaného pufru *wash*
- centrifugovat 2 min./13000g při pokojové teplotě
- odstranit proteklý supernatant, skleněnou fritu nasadit na čistou mikrozkušavku
- přidat 50 µl ddH<sub>2</sub>O, inkubovat 1 min. při pokojové teplotě
- centrifugovat 30 s/7000g při pokojové teplotě
- 1 µl vzorku používat jako templát pro PCR reakce, 35,5 µl vzorku používat pro syntézu homopolymerního konce pomocí terminální deoxynukleotidyltransferázy

#### 4.2.2.1.8 Přečištění ligační směsi pomocí soupravy DNA Clean & Concentrator™-5

- doplnit ligační směs do objemu 100 µl pomocí ddH<sub>2</sub>O, přidat 200 µl dodaného pufru *DNA Binding Buffer*, dokonale promíchat
- přenést do dodané kolonky se skleněnou fritou
- centrifugovat 30 s/14000g při pokojové teplotě
- odstranit proteklý supernatant, přidat 200 µl dodaného pufru *DNA Wash Buffer*
- centrifugovat 30 s/14000g při pokojové teplotě
- odstranit proteklý supernatant, přidat 200 µl dodaného pufru *DNA Wash Buffer*
- centrifugovat 30 s/14000g při pokojové teplotě
- odstranit proteklý supernatant, skleněnou fritu nasadit na čistou mikrozkušavku
- přidat 6–10 µl ddH<sub>2</sub>O, inkubovat 1 min. při pokojové teplotě
- centrifugovat 1 min./14000g při pokojové teplotě

#### 4.2.2.1.9 Přečištění radioaktivně značené DNA pomocí soupravy QIAquick Nucleotide Removal Kit

- přidat 100 µl dodaného pufru *PNI*, dokonale promíchat
- přenést do dodané kolonky se skleněnou fritou
- centrifugovat 1 min./4000g při pokojové teplotě
- odstranit zkumavku s proteklým supernatantem
- kolonku nasadit na čistou zkumavku, přidat 500 µl dodaného pufru *PE*
- centrifugovat 1 min./4000g při pokojové teplotě
- odstranit proteklý supernatant



- centrifugovat 1 min./ 4000g při pokojové teplotě
- odstranit zkumavku s proteklým supernatantem, skleněnou fritu nasadit na čistou mikrozskumavku
- přidat 50 µl dodaného pufru *EB*, inkubovat 1 min. při pokojové teplotě
- centrifugovat 1 min./10000g při pokojové teplotě

#### 4.2.2.1.10 Izolace kvasinkových cytoplasmatických plasmidů a genomové DNA pro reakci PCR z buněk *K. lactis*

- kvasinkovou kulturu nechat narůst 1–2 dny v termostatu při teplotě 28 °C na agarových půdách, YPDA, YPDA + G418 nebo YPDA + HygB
- množství kvasinkové biomasy o objemu ~ 5 µl přenést do mikrozskumavky pipetovacím nástavcem
- přidat 15 µl čerstvě připraveného 20 mM NaOH, dokonale promíchat
- inkubovat 15 min. v termobloku při teplotě 96 °C
- dokonale promíchat
- inkubovat 5 min. při pokojové teplotě
- centrifugovat 5 min./14000g při pokojové teplotě
- použít 1–2 µl supernatantu jako templát pro reakci PCR

#### 4.2.2.1.11 Izolace DNA vázané imobilizovanými proteiny

Tato metoda je vhodná pro testování DNA vazebných aktivit imunoprecipitovaných proteinů, imobilizovaných na kuličky, a byla adaptována podle (Kim 2004).

- k přečištěnému imunoprecipitovanému proteinu přidat 1 ml pufru TGZD
- inkubovat 5 min. na otočném zrcadle při pokojové teplotě
- centrifugovat 5 min./1000g při pokojové teplotě
- odstranit supernatant, k peletu přidat 0,5 ml TGZD + 1 µl LightShift™ Poly (dI-dC) DNA (1 µg/µl) + 10 µl radioaktivně značené DNA (10<sup>4</sup> cpm/µl), připravené podle protokolu 4.2.2.4.7
- inkubovat 30 min. na otočném zrcadle při pokojové teplotě
- centrifugovat 5 min./1000g při pokojové teplotě
- odstranit supernatant, k peletu přidat 1 ml pufru TGZD + 250 mM NaCl, jemně promíchat
- inkubovat 5 min. na otočném zrcadle při pokojové teplotě
- centrifugovat 5 min./1000g při pokojové teplotě
- odstranit supernatant, k peletu přidat 30 µl TE pufru + 1 M NaCl, promíchat
- inkubovat 5 min. při pokojové teplotě
- centrifugovat 5 min./1000g při pokojové teplotě
- supernatant odebrat do čistých mikrozskumavek, přidat k němu polovinu objemu STOP pufru, vzorek DNA analyzovat na akrylamidové elektroforéze

- k peletu přidat 50  $\mu$ l 2x vzorkového pufru pro SDS-PAGE, inkubovat v termobloku 5 min. při teplotě 100 °C, vzorek analyzovat pomocí *western blot*

#### 4.2.2.1.12 Chromatinová imunoprecipitace

- kvasinkovou kulturu nechat narůst přes noc při teplotě 28 °C v 50 ml média YPD, YPD + G418 nebo YPD + HygB
- přeočkovat do 50 ml vytemperovaného média na teplotu 28 °C na OD<sub>600</sub> 0,2
- inkubovat při teplotě 28 °C do požadované optické denzity (OD<sub>600</sub> 4–8)
- přidat k rostoucí kultuře formaldehyd o výsledné koncentraci 1 % (w/v)
- inkubovat 40 min. při teplotě 28 °C
- centrifugovat 10 min./4000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat v 50 ml ddH<sub>2</sub>O
- centrifugovat 10 min./4000g při pokojové teplotě
- umístit na led, dále pracovat při teplotě 4 °C nebo na ledu
- odstranit supernatant, pelet resuspendovat ve 2 ml ledového NDLB pufru
- suspenzi rozdělit do čistých mikrozkušavek po cca 800  $\mu$ l, přidat 400–500  $\mu$ l skleněných kuliček o průměru 0,45 mm
- inkubovat v oscilačním kulovém mlýnku MM301 (Retsch®) 5 min. při frekvenci 30 kmitů/s
- počkat, až se skleněné kuličky samovolně usadí na dně, odebrat suspenzi buněk a pufru do čistých mikrozkušavek
- sonikovat 60x 10 s při 50% amplitudě v sonikátoru Q700 Sonicator (QSONICA), mezi jednotlivými sonikacemi 50 s chladit na ledu
- centrifugovat 5 min./8000g při teplotě 4 °C
- supernatant odebrat do čistých mikrozkušavek
- centrifugovat 20 min./20000g při teplotě 4 °C
- supernatant odebrat do čistých mikrozkušavek, odebrat z něj vzorek (50  $\mu$ l, označení lyzát) a ke zbytku přidat 30  $\mu$ l kuliček anti-HA agarosy HA-7, které byly předtím 3x promyty v NDLB pufru
- inkubovat přes noc na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./1000g při teplotě 4 °C
- odstranit supernatant, k peletu přidat 1 ml vychlazeného ChIP wash pufru s jednovláknovou DNA z lososího mlíčí (100  $\mu$ g/ml), jemně promíchat
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./1000g při teplotě 4 °C
- odstranit supernatant, k peletu přidat 1 ml vychlazeného ChIP wash pufru, jemně promíchat
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./1000g při teplotě 4 °C
- opakovat předchozí tři kroky 1x

- odstranit supernatant, k peletu přidat 1 ml vychlazeného ChIP final wash pufru, jemně promíchat
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./1000g při teplotě 4 °C
- odstranit supernatant, k peletu přidat 120 µl elučního pufru, promíchat
- inkubovat 30 min. při teplotě 37 °C
- supernatant odebrat do čistých mikrozku mávek, přidat k němu 400 µl 1x TBS + 5 µl (20 mg/ml) proteinázy K (totéž udělat se vzorkem lyzát)
- inkubovat 5 hodin při teplotě 65 °C
- přidat 400 µl upraveného fenolu, dokonale promíchat
- centrifugovat 5 min./14000g při pokojové teplotě
- odebrat vrchní vodnou fázi, doplnit její objem do 400 µl TE pufr em a přidat k ní 200 µl upraveného fenolu a 200 µl upraveného chloroformu, dokonale promíchat
- centrifugovat 5 min./14000g při pokojové teplotě
- opakovat předchozí dva kroky 1x
- odebrat vrchní vodnou fázi, doplnit její objem do 400 µl TE pufr em a přidat k ní 400 µl upraveného chloroformu, dokonale promíchat
- centrifugovat 5 min./14000g při pokojové teplotě
- odebrat vrchní vodnou fázi, přidat k ní 0,1 objemu 3 M NaAc, 2,5 objemu 96 % (v/v) ethanolu a 1 µl (25 mg/ml) lineárního polyakrylamidu GenElute™
- inkubovat přes noc při –20 °C
- centrifugovat 10 min./14000g při pokojové teplotě
- odstranit supernatant, přidat 1 ml 70 % (v/v) ethanolu, dokonale promíchat
- centrifugovat 10 min./14000g při pokojové teplotě
- odstranit supernatant, pelet nechat dokonale vyschnout v termostatu při teplotě 37 °C
- rozpustit v 30 µl pufru TE (označení ChIP), analyzovat pomocí PCR či real-time PCR

#### 4.2.2.2 Měření koncentrace DNA

Pro běžnou práci s DNA byla její koncentrace odhadována při elektroforetické analýze. Pro transformaci mikroorganismů a sekvenování byla koncentrace a čistota DNA stanovena spektrofotometricky mikropilárním spektrofotometrem NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies) v objemu 1,5 µl. Pro výpočet koncentrace DNA byla použita hodnota absorbance vzorku při vlnové délce 260 nm a extinkční koeficient 50 µg/ml.

### 4.2.2.3 Elektroforéza DNA

#### 4.2.2.3.1 Agarosová elektroforéza

- agarosu v požadované koncentraci (použité koncentrace od 0,5 % (w/v) do 3 % (w/v) agarosa) rozpustit v 1x pufru TAE
- zchladit na cca 60 °C, nalít do vaničky, přidat EtBr (výsledná koncentrace 0,1 µg/ml) nebo nepřidávat nic v případě barvení DNA pomocí SYBR<sup>TM</sup> Gold, vložit hřeben, nechat agarosu ztuhnout
- vaničku umístit do elektroforetického přístroje, přelít 1x pufrem TAE, vyndat hřeben, nanést vzorky smíchané s 6x vzorkovým pufrem (v případě barvení DNA pomocí SYBR<sup>TM</sup> Gold přidat 1 µl 200x SYBR<sup>TM</sup> Gold) a v případě potřeby nanést standard molekulových hmotností
- zapojit do elektrického pole o vhodném konstantním napětí (použité napětí od 1 V/cm po 8 V/cm) dostatečně dlouhou dobu (použité doby od 20 min. po 48 hodin)
- prohlédnout gel pod transluminátorem generujícím UV záření o vlnové délce 254 nm nebo 312 nm
- v případě potřeby dobarvit gel v roztoku EtBr (výsledná koncentrace 1,6 µg/ml) 2–24 hod. při teplotě 4 °C nebo v 50 ml TE pufru s 1x SYBR<sup>TM</sup> Gold při pokojové teplotě
- vyfotit gel přístrojem Gel Logic 112 Imaging System (Kodak) pod transluminátorem generujícím UV záření o vlnové délce 254 nm
- analyzovat fotografii gelu v programu Carestream (Kodak)

#### 4.2.2.3.2 SDS-agarosová elektroforéza

Tato metoda je vhodná pro dělení DNA s kovalentně asociovanými proteiny a byla adaptována podle (STAM *et al.* 1986).

- agarosu v požadované koncentraci rozpustit v 1x pufru TAE + 0,2 % (w/v) SDS
- zchladit na cca 60 °C, nalít do vaničky, vložit hřeben, nechat agarosu ztuhnout
- vaničku umístit do elektroforetického přístroje, přelít 1x pufrem TAE + 0,2 % (w/v) SDS, vyndat hřeben, nanést vzorky smíchané s 6x vzorkovým pufrem + 0,2 % (w/v) SDS a v případě potřeby nanést standard molekulových hmotností s 0,2 % (w/v) SDS
- zapojit do elektrického pole o vhodném konstantním napětí dostatečně dlouhou dobu
- důkladně opláchnout gel v dH<sub>2</sub>O
- dobarvit gel v roztoku EtBr (výsledná koncentrace 1,6 µg/ml) 2–24 hod. při teplotě 37 °C
- prohlédnout gel pod transluminátorem generujícím UV záření o vlnové délce 254 nm nebo 312 nm
- v případě potřeby vyfotit gel přístrojem Gel Logic 112 Imaging System (Kodak) pod transluminátorem generujícím UV záření o vlnové délce 254 nm

- analyzovat fotografii gelu v programu Carestream (Kodak)

#### 4.2.2.3.3 Polyakrylamidová elektroforéza

- použít skla a další příslušenství pro vektikální elektroforézu
- připravit roztok 7% gelu v 1x TBE pufru (3,5 ml 30 % (w/v) akrylamid + 0,8 % (w/v) N,N'-methylenbisakrylamid; 1,5 ml 10x TBE pufr; 10 ml ddH<sub>2</sub>O; 50 µl 10 % persíran amonný; 10 µl TEMED)
- persíran amonný (APS) a TEMED přidat nakonec, důkladně promíchat
- přenést rychle mezi skla a vložit hřebínky
- vaničku umístit do elektroforetického přístroje, přelít 1x pufrem TBE, vyndat hřeben, nanést vzorky smíchané se STOP pufrem
- zapojit do elektrického pole o vhodném konstantním napětí dostatečně dlouhou dobu
- gel přenést na dvě vrstvy filtračního papíru Whatmann® 3MM, překrýt potravinovou fólií
- sušit 40 min. na sušičce gelů Gel Dryer Model 583 (BIO-RAD) při 80 °C
- založit do kazety s expoziční fólií BAS-MS2040 (Fuji) přes noc
- skenovat expoziční kazetu pomocí Molecular Imager FX (BIO-RAD)
- analyzovat fotografii gelu v programu Quantity One (BIO-RAD)

#### 4.2.2.4 Modifikace DNA

##### 4.2.2.4.1 Štěpení restrikčními endonukleázami

##### Štěpení vektorů a produktů PCR pro přípravu nových rekombinantních vektorů

- 1,5 µl od každého enzymu (koncentrace 10 U/µl) použitého v daném štěpení
- 3 µl nebo 6 µl vhodného pufru
- množství DNA pro štěpení odhadnout z elektroforetické analýzy, pro štěpení použít vektor izolovaný metodou 4.2.2.1.3
- doplnit ddH<sub>2</sub>O do celkového objemu 30 µl, vše dokonale promíchat
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat v termostatu 2–4 hod. při teplotě 37 °C nebo jiné vhodné teplotě
- smíchat s 6x vzorkovým pufrem pro agarosovou elektroforézu a rozdělit elektroforeticky v 0,5 % (w/v) agarosovém gelu
- izolovat naštěpený fragment pomocí metody 4.2.2.1.5

##### Analytické štěpení vektorů a produktů PCR

- 1 µl od každého enzymu použitého v daném štěpení
- 1 µl nebo 2 µl vhodného pufru

- množství DNA pro štěpení odhadnout z elektroforetické analýzy, pro štěpení použít vektor izolovaný metodou 4.2.2.1.2 nebo 4.2.2.1.3
- doplnit ddH<sub>2</sub>O do celkového objemu 10 µl, vše dokonale promíchat
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat v termostatu 1–4 hod. při teplotě 37 °C nebo jiné vhodné teplotě
- smíchat s 6x vzorkovým pufrům pro agarosovou elektroforézu a rozdělit elektroforeticky

#### **Analytické štěpení nových rekombinantních vektorů vzniklých ligací nebo TOPO® TA klonováním**

- 1 µl od každého enzymu použitého v daném štěpení
- 1 µl nebo 2 µl vhodného pufru
- pro štěpení použít 3–5 µl vektoru izolovaného metodou 4.2.2.1.1
- 0,3 µl roztoku TRN
- doplnit ddH<sub>2</sub>O do celkového objemu 10 µl, vše dokonale promíchat
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat v termostatu 2–4 hod. při teplotě 37 °C
- smíchat s 6x vzorkovým pufrům pro agarosovou elektroforézu a rozdělit elektroforeticky

#### **4.2.2.4.2 Defosforylace štěpených vektorů pro přípravu nových rekombinantních vektorů**

- 1 µl SAP fosfatázy
- 2 µl pufru pro SAP fosfatázu
- množství DNA pro defosforylaci odhadnout z elektroforetické analýzy vektoru štěpeného za účelem přípravy nového rekombinantního vektoru
- doplnit ddH<sub>2</sub>O do celkového objemu 20 µl, vše dokonale promíchat
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat v termocykleru 30 min. při teplotě 37 °C
- inkubovat v termocykleru 15 min. při teplotě 65 °C
- inkubovat 5 min. při pokojové teplotě
- pokračovat ligací viz 4.2.2.4.3

#### **4.2.2.4.3 Ligace**

- 2 µl T4 DNA ligázy
- 3 µl pufru pro T4 DNA ligázu
- množství vektoru a insertu pro ligaci odhadnout z elektroforetické analýzy vektoru, při použití defosforylovaného vektoru použít 5 µl insertu
- doplnit ddH<sub>2</sub>O do celkového objemu 30 µl, vše dokonale promíchat
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat v termocykleru 16 hod. při teplotě 16 °C
- inkubovat v termocykleru 10 min. při teplotě 65 °C



- inkubovat 10 min. při teplotě  $-20\text{ }^{\circ}\text{C}$
- použít 5–10  $\mu\text{l}$  pro transformaci kompetentních buněk *E. coli* elektroporací viz 4.2.2.5.1 či ligační směs přechistit metodou 4.2.2.1.8, a poté použít pro transformaci kompetentních buněk *E. coli* elektroporací

#### 4.2.2.4.4 Klonování produktů PCR pomocí soupravy CloneJET PCR Cloning Kit

- PCR produkt připravený pomocí Taq DNA polymerázy izolovat metodou 4.2.2.1.5
- k 5  $\mu\text{l}$  izolovaného PCR produktu přidat 10  $\mu\text{l}$  dodaného 2x reakčního pufru, 2  $\mu\text{l}$  ddH<sub>2</sub>O a 1  $\mu\text{l}$  dodaného *DNA Blunting Enzyme*
- vše dokonale promíchat, centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat v termocykleru 5 min. při teplotě  $70\text{ }^{\circ}\text{C}$
- ochladit na ledu
- přidat 1  $\mu\text{l}$  dodaného vektoru pJET1.2/blunt a 1  $\mu\text{l}$  T4 DNA ligázy
- vše dokonale promíchat, centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat 5 min. při pokojové teplotě
- použít 5–10  $\mu\text{l}$  pro transformaci kompetentních buněk *E. coli* elektroporací viz 4.2.2.5.1 či ligační směs přechistit metodou 4.2.2.1.8, a poté použít pro transformaci kompetentních buněk *E. coli* elektroporací

#### 4.2.2.4.5 Klonování produktů PCR pomocí soupravy TOPO® TA Cloning® Kit

- PCR produkt připravený pomocí Taq DNA polymerázy izolovat metodou 4.2.2.1.5
- k 1–4,75  $\mu\text{l}$  izolovaného PCR produktu přidat 1  $\mu\text{l}$  dodaného 4x ředěného *salt* pufru (300 mM NaCl; 15 mM MgCl<sub>2</sub>) a 0,25  $\mu\text{l}$  pCR™4-TOPO® vektoru
- doplnit ddH<sub>2</sub>O do celkového objemu 6  $\mu\text{l}$ , opatrně promíchat
- inkubovat 5 min. při pokojové teplotě
- ochladit na ledu a přidat 9  $\mu\text{l}$  ddH<sub>2</sub>O
- použít 5–10  $\mu\text{l}$  pro transformaci kompetentních buněk *E. coli* elektroporací viz 4.2.2.5.1 či ligační směs přechistit metodou 4.2.2.1.8, a poté použít pro transformaci kompetentních buněk *E. coli* elektroporací

#### 4.2.2.4.6 Syntéza homopolymerního 3' konce cDNA

- 35,5  $\mu\text{l}$  cDNA izolované metodou 4.2.2.1.7
- 10  $\mu\text{l}$  5x TdT pufru (bez CoCl<sub>2</sub>)
- 2,5  $\mu\text{l}$  10 mM dGTP
- vše dokonale promíchat, inkubovat v termocykleru 3 min. při teplotě  $94\text{ }^{\circ}\text{C}$
- ochladit na ledu, centrifugovat 10 s/14000g při pokojové teplotě
- přidat 2  $\mu\text{l}$  terminální deoxyribonukleotidyltransferázy, dokonale promíchat
- inkubovat v termocykleru 30 min. při teplotě  $37\text{ }^{\circ}\text{C}$
- inkubovat v termocykleru 10 min. při teplotě  $70\text{ }^{\circ}\text{C}$
- použít 2,5  $\mu\text{l}$  pro 5' RACE-PCR

#### 4.2.2.4.7 Fosforylace 5' OH skupin polynukleotidů

- 2–10 µl příslušných polynukleotidů
- 2 µl 10x pufru forward pro T4 polynukleotid kinázu
- 1 µl 1 mM ATP či 3 µl [ $\gamma$  –  $^{32}\text{P}$ ] ATP (10 µCi/µl)
- 2 µl T4 polynukleotid kinázy
- doplnit ddH<sub>2</sub>O do celkového objemu 20 µl, promíchat
- inkubovat 1 hod. při teplotě 37 °C
- v případě radioaktivního značení pokračovat metodou 4.2.2.1.9

#### 4.2.2.5 Vnášení DNA do mikroorganismů

##### 4.2.2.5.1 Transformace buněk *E. coli* elektroporací

- do elektroporační kyvety vychlazené na ledu o elektrodové vzdálenosti 2 mm vnést směs 40–70 µl suspenze elektrokompetentních buněk *E. coli* a 2 µl vhodně naředěné DNA (přibližně 100 ng) nebo 5–10 µl ligační směsi
- inkubovat na ledu 1 min
- aplikovat puls elektroporátorem Gene Pulser Xcell<sup>TM</sup> (BIO-RAD) o kapacitanci 25 µF, napětí 2500 V a odporu 200 Ω
- co nejrychleji po pulzu přidat 1 ml 2xTY média o pokojové teplotě nebo teplotě 37 °C
- přenést do čisté mikrozkušavky
- inkubovat v třepačce pro mikrozkušavky TS-100 (BIOSAN Ltd.) 1 hod./750 otáček za minutu při teplotě 37°C
- transformované buňky vysít v objemech 10 µl, 100 µl a zbylém objemu na agarové půdy s příslušným selekčním antibiotikem, případně ještě přidat X-Gal

##### 4.2.2.5.2 Transformace kvasinek metodou s octanem lithným

- kvasinkovou kulturu nechat narůst 1–2 dny v termostatu při teplotě 28 °C na agarových půdách YPDA
- množství kvasinkové biomasy objemu větší špendlíkové hlavičky (~ 25 µl) přenést pipetovacím nástavcem do mikrozkušavky s 1 ml 100 mM LiAc, dokonale promíchat
- centrifugovat 30 s/14000g při pokojové teplotě
- odstranit supernatant, k peletu přidat v následujícím pořadí: 240 µl PEG, 36 µl 1 M LiAc, 50 µl denaturované jednovláknové DNA z lososího mlíčí (2,0 mg/ml), 25 µl roztoku transformující DNA (100 ng až 5 µg)
- dokonale promíchávat 1 min. při pokojové teplotě
- inkubovat v termobloku 20 min. při teplotě 42 °C
- inkubovat přes noc při pokojové teplotě
- transformované buňky vysít na agarové půdy s příslušným selekčním médiem v objemech 10 µl, 100 µl a zbylém objemu

#### 4.2.2.5.3 Transformace kvasinek metodou s octanem lithným s vysokým výtěžkem

- kvasinkovou kulturu nechat narůst přes noc při teplotě 28 °C v 5 ml média YPD
- přeočkovat do 50 ml vytemperovaného média teploty 28 °C na OD<sub>600</sub> 0,35
- inkubovat při teplotě 28 °C do OD<sub>600</sub> 1
- centrifugovat 10 min./3000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat v 25 ml ddH<sub>2</sub>O
- centrifugovat 10 min./3000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat v 1 ml 100 mM LiAc, přenést do mikrozkušavky
- centrifugovat 30 s/14000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat v 1 ml 100 mM LiAc, rozdělit po 100 µl na jednotlivé transformační reakce do separátních mikrozkušavek
- centrifugovat 30 s/14000g při pokojové teplotě
- odstranit supernatant, k peletu jedné transformační reakce přidat v následujícím pořadí: 240 µl PEG, 36 µl 1 M LiAc, 50 µl denaturované jednovláknové DNA z lososího mlíčí (2,0 mg/ml), 34 µl roztoku transformující DNA (100 ng až 10 µg)
- dokonale promíchávat 1 min. při pokojové teplotě
- inkubovat v termostatu 20 min. při teplotě 28 °C
- inkubovat v termobloku 20 min. při teplotě 42 °C
- centrifugovat 30 s/14000g při pokojové teplotě
- odstranit supernatant, přidat 1 ml sterilní ddH<sub>2</sub>O, dokonale promíchat
- transformované buňky vysít na agarové půdy s příslušným selekčním médiem v objemech 10 µl, 100 µl a zbylém objemu, případně, pokud se jedná o transformaci za účelem tvorby rekombinantních pGKL plasmidů, přenést suspenzi do Erlenmayerovy baňky s 5 ml média YPD
- inkubovat přes noc při 28 °C
- centrifugovat 10 min./3000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat v 1 ml ddH<sub>2</sub>O
- transformované buňky vysít na agarové půdy s příslušným selekčním médiem v objemech 10 µl, 100 µl a zbylém objemu

#### 4.2.2.6 Polymerázová řetězová reakce (PCR)

##### 4.2.2.6.1 Složení směsi pro reakci PCR

- 0,35–1 µl Taq DNA polymerázy či Taq DNA polymerázy FastStart™ (pro analytické reakce PCR a 5' a 3' RACE-PCR), 1 µl Taq DNA polymerázy + 2,5 % (v/v) Pfu DNA polymerázy nebo 1 µl Taq DNA polymerázy + 1-5 % (v/v) Pwo DNA polymerázy (pro PCR reakce sloužící k přípravě rekombinantních vektorů, rekombinantních pGKL plasmidů nebo templátů pro sekvenování)
- 2,5 µl 10x pufru pro Taq DNA polymerázu

- 1,25 µl 10 µM primer 1
- 1,25 µl 10 µM primer 2
- 0,5 µl 10 mM dNTP
- 0,5–4 µl templátové DNA
- doplnit ddH<sub>2</sub>O do celkového objemu 25 µl, dokonale promíchat, vše udržovat na ledu (není potřeba v případě FastStart<sup>TM</sup>)
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat v termocykleru Mastercycler<sup>®</sup> ep Gradient S (Eppendorf) s vhodně zvoleným programem
- 5–10 µl analyzovat elektroforeticky

#### 4.2.2.6.2 Složení směsi pro reakci PCR pomocí FastStart<sup>TM</sup> PCR Master

- 12,5 µl dodaného 2x PCR Master mix
- 1,25 µl 10 µM primer 1
- 1,25 µl 10 µM primer 2
- 1–2,5 µl templátové DNA
- doplnit ddH<sub>2</sub>O do celkového objemu 25 µl, dokonale promíchat
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat v termocykleru Mastercycler<sup>®</sup> ep Gradient S (Eppendorf) s vhodně zvoleným programem
- 5–10 µl analyzovat elektroforeticky

#### 4.2.2.6.3 Program pro analytické PCR reakce a reakce PCR k přípravě rekombinantních vektorů, rekombinantních pGKL plasmidů a templátů pro sekvenování

95 °C	5 min.	} 25 cyklů pro přípravu konstruktů nebo sekvenač. templátů
94 °C	30 s	
variabilní (40–65 °C)	30 s	
72 °C	variabilní (10 s – 3 min. 30 s)	
72 °C	10 min.	
4 °C	do odvolání	} 30-35 cyklů pro účely analýzy

#### 4.2.2.7 Polymerázová řetězová reakce v reálném čase (real-time PCR)

##### 4.2.2.7.1 Složení směsi pro reakci real-time PCR

- 5 µl 2x LightCycler<sup>®</sup> 480 SYBR Green I Master (Roche)
- 1 µl 10 µM primer 1
- 1 µl 10 µM primer 2
- 2,5 µl templátové DNA

- doplnit ddH<sub>2</sub>O do celkového objemu 10 µl, dokonale promíchat, vše udržovat na ledu
- každou reakci připravit v triplicátu
- přenést do plastové destičky LightCycler® 480 Multiwell Plate 384 (Roche)
- zalepit fólií LightCycler® 480 Sealing Foil (Roche)
- centrifugovat 2 min./1500g při pokojové teplotě
- inkubovat v termocykleru LightCycler® 480 Instrument II (Roche) s vhodně zvoleným programem

#### 4.2.2.7.2 Program pro reakci real-time PCR pomocí LightCycler® 480 SYBR Green I

##### Master

95 °C	5 min.	}	45 cyklů
95 °C	10 s		
60 °C	20 s		
72 °C	30 s (měření fluorescence)		
95 °C	5 s	}	měření křivky tání
65 °C	1 min.		
97 °C	variabilní (rychlost změny 0,11 °C/s)		
40 °C	10 s		

#### 4.2.2.8 Sekvenování DNA

##### 4.2.2.8.1 Složení směsi pro sekvenační reakci PCR pomocí soupravy BigDye®

##### Terminator v3.1 Cycle Sequencing Kit

- 2 µl dodané směsi AmpliTaq® DNA polymerázy FS, 2'-deoxyribonukleosid-5'-trifosfátů a fluorescenčních 2'-3'-dideoxyribonukleosid-5'-trifosfátů
- 4 µl dodaného 5x sekvenačního pufru
- 1 µl 5 µM primer (nebo 0,5 µl 10 µM primer)
- 1–8 µl templátové DNA (750 ng v případě vektoru plasmidového typu, 100 ng v případě produktu reakce PCR)
- doplnit ddH<sub>2</sub>O do celkového objemu 20 µl, dokonale promíchat, vše udržovat na ledu
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat v termocykleru Mastercycler® ep Gradient S (Eppendorf) s vhodně zvoleným programem
- pokračovat přečištěním reakce viz 4.2.2.8.3

#### 4.2.2.8.2 Program pro sekvenační reakci PCR pomocí soupravy BigDye® Terminator v3.1 Cycle Sequencing Kit

96 °C	1 min.	}	24 cyklů
96 °C	30 s		
variabilní (50-55 °C)	30 s		
60 °C	4 min.		
4 °C	do odvolání		

#### 4.2.2.8.3 Přečištění sekvenační reakce PCR

- k sekvenační reakci přidat 2 µl 3 M NaAc a 50 µl 96 % (v/v) ethanolu, dokonale promíchat
- inkubovat 30 min. při laboratorní teplotě
- centrifugovat 15 min./13000g při pokojové teplotě
- opatrně odstranit supernatant (pelet není vidět), přidat 250 µl 70 % (v/v) ethanolu, dokonale promíchat
- centrifugovat 15 min./13000g při pokojové teplotě
- důkladně odstranit veškerý supernatant (pelet není vidět)
- inkubovat 10 min. v termostatu při teplotě 37 °C
- předat ke kapilární elektroforetické analýze do Laboratoře sekvenace DNA Přírodovědecké fakulty UK

#### 4.2.2.8.4 Příprava vzorků pro sekvenaci firmou Macrogen

- vektory plasmidového typu určené k sekvenaci izolovat dle protokolu 4.2.2.1.2, naředit pomocí ddH<sub>2</sub>O na výslednou koncentraci 100 ng/µl a 20 µl odeslat k sekvenaci
- produkty reakce PCR naředit pomocí ddH<sub>2</sub>O na výslednou koncentraci 50 ng/µl nebo izolovat dle protokolu 4.2.2.1.5 nebo 4.2.2.1.6 a naředit pomocí ddH<sub>2</sub>O na výslednou koncentraci 50 ng/µl, odeslat k sekvenaci
- pro každých pět vzorků připravit 20 µl 10 µM primeru v ddH<sub>2</sub>O, odeslat spolu s příslušnými vzorky k sekvenaci

#### 4.2.2.8.5 Příprava vzorků pro sekvenaci Laboratoří sekvenace DNA PřF UK

- vektory plasmidového typu určené k sekvenaci izolovat dle protokolu 4.2.2.1.2, do reakce dát množství DNA odpovídající 3 ng/100 párů bází délky vektoru, objem vzorku nesmí přesáhnout 7,5 µl
- produkty reakce PCR izolovat dle protokolu 4.2.2.1.5 nebo 4.2.2.1.6, do reakce dát množství DNA odpovídající 5 ng/100 párů bází délky produktu PCR, objem vzorku nesmí přesáhnout 7,5 µl
- 0,5 µl 10 µM primer
- doplnit ddH<sub>2</sub>O do celkového objemu 8 µl, dokonale promíchat, odeslat k sekvenaci



### 4.2.3 Práce s RNA

#### 4.2.3.1 Izolace RNA

##### 4.2.3.1.1 Izolace celkové kvasinkové RNA pomocí horkého kyselého fenolu

- kvasinkovou kulturu nechat narůst přes noc při teplotě 28 °C v 50 ml média YPD, YPD + G418 nebo YPD + HygB
- přeočkovat do 25 ml vytemperovaného média teploty 28 °C na OD<sub>600</sub> 0,1–0,2
- inkubovat při teplotě 28 °C do požadované optické denzity (OD<sub>600</sub> 1–2)
- ke kultuře přidat 25 ml ledu a centrifugovat 5 min./4000g při teplotě 4 °C
- odstranit supernatant, pelet resuspendovat v 8 ml izolačního roztoku a neprodleně přidat 12 ml kyselého fenolu temperovaného na teplotu 65 °C
- inkubovat 5 min. ve vodní lázni o teplotě 65 °C, během toho 3x 30 s promíchávat
- centrifugovat 5 min./4000g při pokojové teplotě
- odstranit spodní organickou fázi, přidat 8 ml kyselého fenolu o teplotě 65 °C
- inkubovat 5 min. ve vodní lázni o teplotě 65 °C, během toho 3x 30 s promíchávat
- centrifugovat 5 min./4000g při pokojové teplotě
- odstranit spodní organickou fázi, přidat 6 ml kyselého fenolu a 6 ml upraveného chloroformu, 30 s promíchávat
- centrifugovat 5 min./4000g při pokojové teplotě
- horní vodnou fázi přenést do čisté zkumavky, přidat 5,5 ml upraveného chloroformu, 30 s promíchávat
- centrifugovat 2 min./4000g při pokojové teplotě
- odebrat vrchní vodnou fázi a rozdělit do mikrozkušavek po objemech 400 µl, přidat k ní 0,1 objemu 3 M NaAc a 2,5 objemu 96 % (v/v) ethanolu
- inkubovat 1 hod. či přes noc při –20 °C
- centrifugovat 30 min./14000g při teplotě 4 °C
- odstranit supernatant, přidat 1 ml 70 % (v/v) ethanolu vychlazeného na –20 °C, dokonale promíchat
- centrifugovat 10 min./14000g při teplotě 4 °C
- odstranit supernatant, pelet nechat dokonale vyschnout na ledu
- rozpustit v 10–20 µl ddH<sub>2</sub>O ošetřené DEPC
- 1–3 µl vzorku analyzovat elektroforeticky

##### 4.2.3.1.2 Izolace celkové kvasinkové RNA pomocí TRI Reagent® Solution

- kvasinkovou kulturu nechat narůst přes noc při teplotě 28 °C v 50 ml média YPD, YPD + G418 nebo YPD + HygB
- přeočkovat do 25 ml vytemperovaného média teploty 28 °C na OD<sub>600</sub> 0,1–0,2
- inkubovat při teplotě 28 °C do požadované optické denzity (OD<sub>600</sub> 1–2)
- centrifugovat 5 min./4000g při teplotě 4 °C

- odstranit supernatant, pelet resuspendovat v 1 ml TRI Reagent<sup>®</sup> Solution
- inkubovat 5 min. při pokojové teplotě
- centrifugovat 10 min./12000g při teplotě 4 °C
- supernatant převést do čisté mikrozkušavky, přidat 200 µl chloroformu (bez isoamylalkoholu), promíchat
- inkubovat 5–15 min. při pokojové teplotě
- centrifugovat 15 min./12000g při teplotě 4 °C
- horní vodnou fázi přenést do čisté zkumavky, přidat 500 µl isopropanolu, 10 s promíchávat
- inkubovat 5–10 min. při pokojové teplotě
- centrifugovat 8 min./12000g při teplotě 4 °C
- odstranit supernatant, přidat 1 ml 75 % ethanolu
- centrifugovat 5 min./7500g při teplotě 4 °C
- odstranit supernatant, pelet nechat dokonale vyschnout na ledu
- rozpustit v 10–20 µl ddH<sub>2</sub>O ošetřené DEPC
- 1–3 µl vzorku analyzovat elektroforeticky

#### 4.2.3.1.3 Izolace celkové RNA savčích tkáňových kultur pomocí RNA Blue

- kulturu adherentních tkáňových buněk na Petriho misce (Ø 60 či 34 mm) postavit na led a odsát médium
- přidat 1 ml RNA Blue, rozvolnit buňky pipetováním, přenést do centrifugační zkumavky
- inkubovat 5 min. na ledu
- přidat 200 µl chloroformu, 15 s promíchávat
- inkubovat 5 min. při pokojové teplotě
- centrifugovat 10 min./12000g při teplotě 4 °C
- horní vodnou fázi přenést do čisté zkumavky, přidat 500 µl isopropanolu, 10 s promíchávat
- inkubovat 10 min. na ledu
- centrifugovat 10 min./12000g při teplotě 4 °C
- odstranit supernatant, přidat 1 ml 75 % ethanolu
- centrifugovat 5 min./12000g při teplotě 4 °C
- odstranit supernatant, pelet nechat dokonale vyschnout na ledu
- rozpustit v 50 µl ddH<sub>2</sub>O ošetřené DEPC
- 2 µl vzorku analyzovat elektroforeticky

#### 4.2.3.1.4 Degradace DNA ve vzorcích izolované RNA pomocí deoxyribonukleázy I

- 10–44 µl vzorku RNA
- 5 µl 10x pufr pro DNázu I
- 1 µl DNázy I
- doplnit ddH<sub>2</sub>O ošetřenou DEPC do celkového objemu 50 µl, dokonale promíchat

- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat 40 min. při teplotě 37 °C
- přidat 10 µl inaktivačního reagensu dodaného v soupravě DNA-free™ Kit
- inkubovat 2 min. při pokojové teplotě
- centrifugovat 1,5 min./10000g při pokojové teplotě
- supernatant odebrat do čisté zkumavky, skladovat v –80 °C

#### 4.2.3.1.5 Zachycení srážené RNA na nitrocelulózový filtr

- ke vzorku transkripční reakce *in vitro* přidat 20 µl BSA (5 mg/ml)
- odebrat do čisté zkumavky, přidat 1,2 ml 12 % (v/v) TCA vychlazené na ledu
- inkubovat 10 min. na ledu, promíchat
- přefiltrovat vzorek na skleněné fritě přes nitrocelulosový filtr o poréznosti 0,2 µm
- promýt filtr cca 5 ml 5 % (v/v) TCA + 0,1 % (w/v) pyrofosforečnanem sodným
- umístit filtr do scintilační zkumavky a zalít scintilačním roztokem EcoLume™ Liquid Scintillation Cocktail (ICN)

#### 4.2.3.2 Měření koncentrace RNA

Pro veškerou práci s RNA byla koncentrace a čistota RNA stanovena spektrofotometricky mikrokapilárním spektrofotometrem NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies) v objemu 1,5 µl. Pro výpočet koncentrace DNA byla použita hodnota absorpance vzorku při vlnové délce 260 nm a extinkční koeficient 40 µg/ml.

#### 4.2.3.3 Elektroforéza RNA

##### 4.2.3.3.1 Nedenaturační agarosová elektroforéza

- agarosu v požadované koncentraci (použité koncentrace 1,2–1,5 % (w/v) agarosa) rozpustit v 1x pufru TAE
- zchladit na cca 60 °C, nalít do vaničky, přidat EtBr (výsledná koncentrace 0,1 µg/ml), vložit hřeben, nechat agarosu ztuhnout
- vaničku umístit do elektroforetického přístroje, přelít 1x pufrem TAE, vyndat hřeben, nanést vzorky smíchané s 6x vzorkovým pufrem a v případě potřeby nanést standard molekulových hmotností
- zapojit do elektrického pole o vhodném konstantním napětí (použité napětí od 5 V/cm po 8 V/cm) dostatečně dlouhou dobu (použité doby od 50 min. po 1,5 hodiny)
- prohlédnout gel pod transluminátorem generujícím UV záření o vlnové délce 254 nm nebo 312 nm
- vyfotit gel přístrojem Gel Logic 112 Imaging System (Kodak) pod transluminátorem generujícím UV záření o vlnové délce 254 nm

- analyzovat fotografii gelu v programu Carestream (Kodak)

#### 4.2.3.3.2 Denaturační agarosová elektroforéza

Tato metoda byla adaptována podle (MAŠEK *et al.* 2005).

- postupovat stejně jako při nedenaturační agarosové elektroforéze RNA
- ke vzorku s 6x vzorkovým pufrem přidat formamid do celkové koncentrace 60 % (v/v)
- inkubovat v termobloku 5 min. při teplotě 65 °C
- ochladit vzorky před nanesením na gel inkubací 5 min. na ledu

#### 4.2.3.4 Modifikace RNA

##### 4.2.3.4.1 Oligocapping pomocí soupravy FirstChoice® RLM-RACE Kit

Tato metoda slouží k analýze eukaryotických transkriptů obsahujících methylguanozinovou čepičku pomocí 5' RACE-PCR. Všechny enzymy a jejich pufrů byly součástí soupravy.

- použít 1–10 µg celkové RNA
- 2 µl 10x pufru pro alkalickou fosfatázu z telecího střeva (CIP)
- 2 µl CIP
- doplnit ddH<sub>2</sub>O ošetřenou DEPC do celkového objemu 20 µl, dokonale promíchat
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat 1 hod. při teplotě 37 °C
- přidat 15 µl dodaného roztoku octanu amonného, 115 µl ddH<sub>2</sub>O ošetřené DEPC, 75 µl kyselého fenolu, 75 µl upraveného chloroformu, dokonale promíchat
- centrifugovat 5 min./14000g při teplotě 4 °C
- odebrat vrchní vodnou fázi do čisté zkumavky, přidat 150 µl upraveného chloroformu, dokonale promíchat
- centrifugovat 5 min./14000g při teplotě 4 °C
- odebrat vrchní vodnou fázi do čisté zkumavky, přidat 150 µl vychlazeného isopropanolu, dokonale promíchat, inkubovat 10 min. na ledu
- centrifugovat 20 min./19000g při teplotě 4 °C
- odstranit supernatant, přidat 0,5 ml vychlazeného 70 % ethanolu
- centrifugovat 10 min./19000g při teplotě 4 °C
- odstranit supernatant, pelet nechat dokonale vyschnout na ledu
- rozpustit v 6 µl 1x pufru pro kyselou pyrofosfatázu z tabáku (TAP)
- 3 µl použít na reakci s 1 µl TAP, 3 µl použít na slepou reakci s 1 µl ddH<sub>2</sub>O ošetřené DEPC, dokonale promíchat
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat 1 hod. při teplotě 37 °C
- k reakcím přidat 0,5 µl 5' RACE Adapter, 0,75 µl 10x pufru pro T4 RNA ligázu, 1 µl T4 RNA ligázy (2,5 U/µl), 1,25 µl ddH<sub>2</sub>O ošetřené DEPC, dokonale promíchat

- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat 1 hod. při teplotě 37 °C
- pokračovat reverzní transkripcí podle protokolu 4.2.3.6

#### 4.2.3.4.2 Syntéza homopolymerního 3' konce RNA

- použít 1–10 µg celkové RNA
- 5 µl 5x pufru pro poly(A) polymerázu z *Escherichia coli* (E-PAP)
- 2,5 µl 25 mM MnCl<sub>2</sub>
- 2 µl E-PAP (2 U/µl)
- 1,25 µl 40 mM CTP
- 1 µl ribonukleázových inhibitorů či 0,5 µl RiboLock™
- doplnit ddH<sub>2</sub>O ošetřenou DEPC do celkového objemu 25 µl, promíchat
- centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat 90 min. při teplotě 37 °C
- 10 µl použít pro reverzní transkripci s primerem oligo(dG)anch2 podle protokolu 4.2.3.6

#### 4.2.3.5 Transkripční reakce *in vitro*

Tato metoda byla použita k analýze RNA polymerázové aktivity afinitně izolované RNA polymerázy kvasinkových lineárních plasmidů pomocí GFP-Trap®\_A.

- použít 1–5 µg templátové DNA
- 10 µl 10x transkripční pufr
- 10 µl 0,5 M KCl
- 1 µl 100 mM MnCl<sub>2</sub>
- 10 µl BSA (1 mg/ml)
- 10 µl 10 mM ATP, CTP, GTP
- 5–10 µl 0,2 mM UTP
- 0,2–0,28 µl [ $\alpha$  - <sup>32</sup>P] UTP (3000 Ci/mmol, 10 µCi/µl)
- 1 µl RiboLock™
- doplnit ddH<sub>2</sub>O ošetřenou DEPC do celkového objemu 100 µl, promíchat
- odebrat 1–2 µl jako kontrolu celkového množství radioaktivity v reakci a umístit do scintilační zkumavky se scintilačním roztokem EcoLume™ Liquid Scintillation Cocktail (MP Biomedicals)
- inkubovat v hybridizační peci 30–60 min. při teplotě 30 °C za občasných míchání
- inkubovat 10 min. v termobloku při teplotě 100 °C
- pokračovat srážením RNA podle protokolu 4.2.3.1.5

#### 4.2.3.6 Reverzní transkripce

##### 4.2.3.6.1 Složení směsi pro reverzní transkripci

- použít 1–3 µg celkové RNA
- 0,5 µl RiboLock™
- 1,5 µl náhodných hexanukleotidových primerů (150 ng/µl) či 1 µl oligod(T)anchor nebo oligo(dG)anch2 primeru (1 µg/µl)
- 1 µl 10 mM dNTP
- doplnit ddH<sub>2</sub>O ošetřenou DEPC do celkového objemu 14,5 µl, dokonale promíchat, vše udržovat na ledu
- inkubovat 3–5 min. v termobloku při teplotě 65 °C, ochladit na ledu
- centrifugovat 10 s/14000g při teplotě 4 °C
- přidat 4 µl 5x pufru pro SuperScript III™ reverzní transkriptázu (SSCIII)
  - 1 µl DTT (0,1 M)
  - 0,5 µl SSCIII (200 U/µl)
- dokonale promíchat, centrifugovat 10 s/14000g při pokojové teplotě
- inkubovat v termocykleru Mastercycler® ep Gradient S (Eppendorf) s vhodně zvoleným programem
- komplementární DNA izolovat podle protokolu 4.2.2.1.7

##### 4.2.3.6.2 Program pro reverzní transkripci s náhodnými primery

25 °C	10 min.
50 °C	99 min.
70 °C	15 min.
4 °C	do odvolání

##### 4.2.3.6.3 Program pro reverzní transkripci s oligod(T)anchor nebo oligo(dG)anch2 primerem

55 °C	75 min.
70 °C	15 min.
4 °C	do odvolání

#### 4.2.4 Práce s proteiny

##### 4.2.4.1 Příprava bakteriálních lyzátů

- bakteriální buňky propláchnuté puforem TKG po indukci IPTG, viz protokol 4.2.1.5, dát na led a od tohoto kroku udržovat na ledu nebo při 4 °C



- resuspendovat buňky v roztoku B-PER<sup>®</sup> Bacterial Protein Extraction Reagent v objemu 1 ml na 50 ml vstupní kultury, přidat směs inhibitorů proteáz (1 tableta směsi inhibitorů Roche na 10 ml roztoku B-PER<sup>®</sup>) a 1 mM PMSF
- inkubovat 1 hod. na ledu
- sonikovat 30x 30 s při 60% amplitudě v sonikátoru Q700 Sonicator (QSONICA), mezi jednotlivými sonikacemi chladit 1 min. na ledu
- centrifugovat 20 min./20000g při teplotě 4 °C
- odstranit pelet, supernatant přenést do čisté mikrozkušavky a uchovávat v –80 °C
- část supernatantu analyzovat pomocí SDS-PAGE proteinové elektroforézy
- pokračovat protokolem 4.2.4.2

#### 4.2.4.2 Purifikace proteinů s GST-epitopem z bakteriálních a kvasinkových lyzátů

- jemně promíchat zásobní suspenzi Glutathione Sepharose<sup>™</sup> 4 Fast Flow uchovávanou při teplotě 4 °C, odebrat objem odpovídající 1/100 objemu přečištěného lyzátu
- centrifugovat 5 min./500g při teplotě 4 °C
- odstranit supernatant, přidat 1 ml TKG vychlazeného na ledu, jemně promíchat
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- opakovat předchozí tři kroky 2-3x
- odstranit supernatant, k peletu přidat přečištěný lyzát
- inkubovat minimálně 1 hod. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- ze supernatantu odebrat vzorek na SDS-PAGE proteinovou elektroforézu (označení *flow*), zbytek odstranit, k peletu přidat 1 ml TKG vychlazeného na ledu
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- odstranit supernatant (možno odebrat vzorek na SDS-PAGE proteinovou elektroforézu, označení *wash* 1), k peletu přidat 1 ml TKG vychlazeného na ledu
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- opakovat předchozí tři kroky 1–2x (označení volitelně odebíraných vzorků na SDS-PAGE proteinovou elektroforézu *wash* 2-4)
- odstranit supernatant, k peletu přidat 1 objem elučního pufru 1
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- odebrat supernatant do čisté mikrozkušavky, odebrat z něj vzorek na SDS-PAGE proteinovou elektroforézu (označení *eluce* 1), k peletu přidat 1 objem elučního pufru 2
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C

- odebrat supernatant do čisté mikrozkušavky, odebrat z něj vzorek na SDS-PAGE proteinovou elektroforézu (označení *eluce* 2)
- ke zbylému peletu přidat 1 objem 2x vzorkového pufru pro SDS-PAGE proteinovou elektroforézu a odebrat vzorek na SDS-PAGE proteinovou elektroforézu (označení *matrice*)

#### 4.2.4.3 Purifikace proteinů s HIS-epitopem z bakteriálních a kvasinkových lyzátů

- jemně promíchat zásobní suspenzi Ni-NTA Superflow uchovávanou při teplotě 4 °C, odebrat objem odpovídající 400 µl zásobní suspenze na 50 ml vstupní kultury
- centrifugovat 5 min./500g při teplotě 4 °C
- odstranit supernatant, přidat 1 ml TN vychlazeného na ledu, jemně promíchat
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- opakovat předchozí tři kroky 2-3x
- odstranit supernatant, k peletu přidat přechištěný buněčný lyzát
- inkubovat minimálně 1 hod. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- ze supernatantu odebrat vzorek na SDS-PAGE proteinovou elektroforézu (označení *flow*), zbytek odstranit, k peletu přidat 1 ml TN vychlazeného na ledu
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- odstranit supernatant (možno odebrat vzorek na SDS-PAGE proteinovou elektroforézu, označení *wash* 1), k peletu přidat 1 ml TN vychlazeného na ledu
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- opakovat předchozí tři kroky 1–2x s omývacím pufrem TN + 10 mM imidazol a 2 mM β-merkapt ethanol (označení volitelně odebíraných vzorků na SDS-PAGE proteinovou elektroforézu *wash* 2-4)
- odstranit supernatant, k peletu přidat 1 objem elučního pufru 1
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- odebrat supernatant do čisté mikrozkušavky, odebrat z něj vzorek na SDS-PAGE proteinovou elektroforézu (označení *eluce* 1), k peletu přidat 1 objem elučního pufru 2
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- odebrat supernatant do čisté mikrozkušavky, odebrat z něj vzorek na SDS-PAGE proteinovou elektroforézu (označení *eluce* 2)
- ke zbylému peletu přidat 1 objem 2x vzorkového pufru pro SDS-PAGE proteinovou elektroforézu a odebrat vzorek na SDS-PAGE proteinovou elektroforézu (označení *matrice*)

#### 4.2.4.4 Příprava kvasinkových lyzátů

##### 4.2.4.4.1 Příprava kvasinkových lyzátů s vysokým výtěžkem

- kvasinkovou kulturu nechat narůst přes noc při teplotě 28 °C v 5 ml média YPD, YPD + G418 nebo YPD + HygB
- přeočkovat do 5 ml vytemperovaného média teploty 28 °C na OD<sub>600</sub> 0,05–0,2
- inkubovat při teplotě 28 °C do požadované optické denzity (OD<sub>600</sub> 1–2)
- 3 ml kultury zchladit na ledu, od tohoto kroku vše při 4 °C nebo na ledu
- centrifugovat 30 s/14000g při teplotě 4 °C
- odstranit supernatant, k peletu přidat 1 ml ddH<sub>2</sub>O a důkladně promíchat
- centrifugovat 30 s/14000g při teplotě 4 °C
- odstranit supernatant, pelet zmrazit v kapalném dusíku 3 min.
- k peletu přidat 100 µl 1 M NaOH + 7 % (v/v) β-merkapt ethanol (čerstvě připravit)
- inkubovat 2 min. při pokojové teplotě, průběžně promíchat
- přidat 100 µl 50 % (v/v) TCA
- inkubovat 5 min. na ledu
- centrifugovat 5 min./14000g při teplotě 4 °C
- dokonale odstranit supernatant, pelet resuspendovat pipetováním v 500 µl ledového roztoku 1 M Tris-HCl pH 8
- centrifugovat 1 min./14000g při teplotě 4 °C
- odstranit supernatant, k peletu přidat 100–150 µl 2x vzorkového pufru pro SDS-PAGE proteinovou elektroforézu, dokonale promíchat
- inkubovat 5 min. v termobloku při teplotě 100 °C
- 5–15 µl analyzovat pomocí SDS-PAGE proteinové elektroforézy

##### 4.2.4.4.2 Příprava kvasinkových lyzátů z protoplastů

- kvasinkovou kulturu nechat narůst přes noc při teplotě 28 °C v 50 ml média YPD, YPD + G418 nebo YPD + HygB
- přeočkovat do 50–500 ml vytemperovaného média teploty 28 °C na OD<sub>600</sub> 0,2
- inkubovat při teplotě 28 °C do požadované optické denzity (OD<sub>600</sub> 4–8)
- centrifugovat 10 min./4000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat v čerstvě připravených 30 ml 60 mM EDTA-NaOH pH 8 + 1 % (v/v) β-merkapt ethanol
- inkubovat 20 min. při pokojové teplotě
- centrifugovat 5 min./4000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat v 50 ml pufru TKCl
- centrifugovat 5 min./4000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat v 30 ml pufru TKCl + 5 mg lytikázy
- inkubovat 10 min. při pokojové teplotě

- (kontrola protoplastů: 20  $\mu$ l suspenze kápnout na podložní sklo, přikrýt krycím sklem, pozorovat v mikroskopu při zvětšení 500x, ke hraně krycího skla přidat 10–20  $\mu$ l dH<sub>2</sub>O, protoplasty praskají)
- centrifugovat 5 min./3000g při pokojové teplotě
- odstranit supernatant, pelet opatrně kývavě otáčivým pohybem resuspendovat v 50 ml pufru TKCl
- centrifugovat 5 min./3000g při pokojové teplotě
- odstranit supernatant, pelet opatrně kývavě otáčivým pohybem resuspendovat v 50 ml pufru TKCl
- centrifugovat 5 min./3000g při pokojové teplotě
- dokonale odstranit veškerý supernatant, pelet zchladit na ledu 5 min.
- pelet opatrně kývavě otáčivým pohybem resuspendovat v 2–7 ml ledového pufru TM + 1 tableta směsi inhibitorů proteáz a 1 mM PMSF
- přenést pelet do Dounceova homogenizátoru vychlazeného na ledu, skleněným pístem pomalu pohybovat 5x nahoru a dolů
- (kontrola lyze protoplastů: 20  $\mu$ l suspenze kápnout na podložní sklo, přikrýt krycím sklem, pozorovat v mikroskopu při zvětšení 500x)
- centrifugovat 20 min./20000g při 4 °C
- supernatant odebrat do čistých mikrozkušavek, možno uchovávat při teplotě –80 °C
- 10–25  $\mu$ l analyzovat pomocí SDS-PAGE proteinové elektroforézy

#### 4.2.4.4.3 Příprava kvasinkových lyzátů pomocí skleněných kuliček

- kvasinkovou kulturu nechat narůst přes noc při teplotě 28 °C v 50 ml média YPD, YPD + G418 nebo YPD + HygB
- přeočkovat do 100–1000 ml vytemperovaného média teploty 28 °C na OD<sub>600</sub> 0,2
- inkubovat při teplotě 28 °C do požadované optické denzity (OD<sub>600</sub> 4–8)
- centrifugovat 10 min./4000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat v 50 ml ddH<sub>2</sub>O
- centrifugovat 10 min./4000g při pokojové teplotě
- umístit na led, dále pracovat při teplotě 4 °C nebo na ledu
- odstranit supernatant, pelet resuspendovat ve 2–3 ml ledového *GFP wash* pufru na 1 ml buněk
- suspenzi rozdělit do čistých mikrozkušavek po cca 800  $\mu$ l, přidat 400–500  $\mu$ l skleněných kuliček o průměru 0,45 mm
- inkubovat v oscilačním kulovém mlýnku MM301 (Retsch®) 5 min. při frekvenci 30 kmitů/s
- počkat, až se skleněné kuličky samovolně usadí na dně, odebrat supenzi buněk a pufru do čistých mikrozkušavek
- centrifugovat 5 min./8000g při teplotě 4 °C
- supernatant odebrat do čistých mikrozkušavek
- inkubovat 10 min. na ledu

- centrifugovat 20 min./20000g při teplotě 4 °C
- odebrat supernatant do čisté mikrozkušavky, odebrat z něj vzorek na SDS-PAGE proteinovou elektroforézu
- 15–25 µl analyzovat pomocí SDS-PAGE proteinové elektroforézy

#### 4.2.4.5 Imunoprecipitace proteinů z kvasinek

- kvasinkovou kulturu nechat narůst přes noc při teplotě 28 °C v 50 ml média YPD, YPD + G418 nebo YPD + HygB
- přeočkovat do 50–500 ml vytemperovaného média teploty 28 °C na OD<sub>600</sub> 0,2
- inkubovat při teplotě 28 °C do požadované optické denzity (OD<sub>600</sub> 4–8)
- buňky umístit na led, dále pracovat při teplotě 4 °C nebo na ledu
- centrifugovat 10 min./4000g při teplotě 4 °C
- odstranit supernatant, pelet resuspendovat v 40 ml ddH<sub>2</sub>O vychlazené na ledu
- centrifugovat 10 min./4000g při teplotě 4 °C
- odstranit supernatant, pelet resuspendovat v 1 objemu vychlazené ddH<sub>2</sub>O, suspenzi rozdělit do čistých mikrozkušavek po 800 µl
- centrifugovat 30 s/14000g při teplotě 4 °C
- odstranit supernatant, k peletu v každé mikrozkušavce přidat 1,5 objemu skleněných kuliček o průměru 0,45 mm a 400 µl vychlazeného nedenaturujícího lyzačního pufru + směs inhibitorů proteáz + 1 mM PMSF, dokonale promíchat
- inkubovat v oscilačním kulovém mlýnku MM301 (Retsch®) 5 min. při frekvenci 30 kmitů/s (alternativně intenzivně promíchat 5x 30 s v chladové místnosti, mezitím inkubovat na ledu)
- počkat, až se skleněné kuličky samovolně usadí na dně, odebrat supenzi buněk a pufru do čistých mikrozkušavek
- centrifugovat 5 min./8000g při teplotě 4 °C
- supernatant odebrat do čistých mikrozkušavek
- inkubovat 10 min. na ledu
- centrifugovat 20 min./20000g při teplotě 4 °C
- supernatant odebrat do čistých mikrozkušavek, možno uchovávat při teplotě –80 °C, 15–25 µl analyzovat pomocí SDS-PAGE proteinové elektroforézy (označení lyzát)
- zásobní suspenzi Protein G-agarosy uchovávanou při teplotě 4 °C jemně promíchat a odebrat 20–30 µl zásobní suspenze na každých 50 ml vstupní kultury
- centrifugovat 5 min./500g při teplotě 4 °C
- odstranit supernatant, k Protein G-agarose přidat 600 µl vychlazeného nedenaturujícího lyzačního pufru + směs inhibitorů proteáz + 1 mM PMSF, jemně promíchat
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- opakovat předchozí tři kroky 2x

- odstranit supernatant, k Protein G-agarose přidat 1 ml vychlazeného nedenaturujícího lyzačního pufru + směs inhibitorů proteáz + 1 mM PMSF, jemně promíchat, odebrat 75  $\mu$ l a přidat k lyzátu
- inkubovat minimálně 1 hod. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- odebrat supernatant do čisté mikrozkušavky, přidat 925  $\mu$ l promyté Protein G-agarosy, přidat protilátku v požadovaném množství (většinou 1-2  $\mu$ l)
- inkubovat přes noc na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- odebrat supernatant (volitelně analyzovat 15–25  $\mu$ l pomocí SDS-PAGE proteinové elektroforézy – označení *flow*), k peletu přidat 4 ml vychlazeného nedenaturujícího lyzačního pufru + směs inhibitorů proteáz + 1 mM PMSF, jemně promíchat
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- odstranit supernatant, k peletu přidat 4 ml vychlazeného nedenaturujícího lyzačního pufru + směs inhibitorů proteáz + 1 mM PMSF, jemně promíchat
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./500g při teplotě 4 °C
- opakovat předchozí tři kroky 1x
- k peletu přidat stejný objem 2x vzorkového pufru pro SDS-PAGE proteinovou elektroforézu, 20–25  $\mu$ l analyzovat pomocí SDS-PAGE proteinové elektroforézy (označení IP = imunoprecipitace)

#### 4.2.4.6 Imunoprecipitace proteinů s GFP-epitopem z kvasinkových lyzátů

- lyzát požadované kvasinkové kultury připravit podle protokolu 4.2.4.4.2 nebo podle protokolu 4.2.4.4.3
- důkladně promíchat zásobní suspenzi GFP-Trap®\_A uchovávanou při 4 °C, odebrat 20–35  $\mu$ l do čisté mikrozkušavky, přidat 1 ml vychlazeného *GFP wash* pufru + směs inhibitorů proteáz + 1 mM PMSF, důkladně promíchat
- centrifugovat 5 min./1000g při teplotě 4 °C
- odstranit supernatant, k GFP-Trap®\_A přidat 1 ml vychlazeného *GFP wash*, jemně promíchat
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./1000g při teplotě 4 °C
- opakovat předchozí tři kroky 2x
- odstranit supernatant, přidat lyzát
- inkubovat 24 hod. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./1000g při teplotě 4 °C
- odebrat supernatant (volitelně analyzovat 15–25  $\mu$ l pomocí SDS-PAGE proteinové elektroforézy – označení *flow*), k peletu přidat 1 ml vychlazeného *GFP wash* + směs inhibitorů proteáz + 1 mM PMSF, jemně promíchat



- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./1000g při teplotě 4 °C
- odstranit supernatant, k peletu přidat 1 ml vychlazeného *GFP wash* + směs inhibitorů proteáz + 1 mM PMSF, jemně promíchat
- inkubovat 10 min. na otočném zrcadle při teplotě 4 °C
- centrifugovat 5 min./1000g při teplotě 4 °C
- opakovat předchozí tři kroky 1x
- k peletu přidat stejný objem 2x vzorkového pufru pro SDS-PAGE proteinovou elektroforézu, 20–25 µl analyzovat pomocí SDS-PAGE proteinové elektroforézy (označení IP = imunoprecipitace)

#### 4.2.4.7 Imunoprecipitace proteinů s GST-epitopem z kvasinkových lyzátů

- postupovat stejně jako při protokolu 4.2.4.6
- místo GFP-Trap<sup>®</sup>\_A použít GST-Trap<sup>®</sup>\_A

#### 4.2.4.8 Polyakrylamidová proteinová elektroforéza SDS-PAGE

##### 4.2.4.8.1 Příprava polyakrylamidového gelu

- použít skla a další příslušenství pro elektroforézu ze soupravy Mini-PROTEAN<sup>®</sup> 3 (BIO-RAD)
- podle hodnot uvedených v Tabulce 10 připravit roztok dělicího gelu, persíran amonný (APS) a TEMED přidat nakonec, důkladně promíchat
- přenést rychle mezi skla, převrstvit převrstvovacím pufrem
- inkubovat cca 30 min. při pokojové teplotě
- odstranit převrstvovací pufr a vysušit skla nad gelem filtračním papírem
- připravit roztok zaostřovacího gelu, složení na 2 gely: 0,65 ml směsi akrylamidu + N,N'-methylenbisakrylamidu; 1,25 ml 4x Tris-HCl/SDS pH 6,8; 3,05 ml ddH<sub>2</sub>O; nakonec přidat 37,5 µl APS a 7,5 µl TEMED
- přelít zaostřovacím gelem a vložit hřebínky
- inkubovat cca 20 min. při pokojové teplotě
- gely je možno uchovávat několik dní zabalené ve filtračním papíru namočeném v převrstvovacím pufru nebo ddH<sub>2</sub>O a neprodyšně uzavřené při teplotě 4 °C

##### 4.2.4.8.2 Příprava gradientového polyakrylamidového gelu

- připravit dva roztoky dělicího gelu pro krajní hodnoty koncentrace požadovaného gradientu v objemu 2,75–3 ml (objemy na 1 gradientový gel) se zachováním poměru hodnot uvedených v Tabulce 10
- přenést oba roztoky do míchačky gradientů, roztok o vyšší koncentraci umístit do přihrádky blíže k výpusti

- vložit do roztoku o vyšší koncentraci magnetické míchátko, zapnout míchání, otevřít uzávěr mezi roztoky a výpusť umístěnou mezi skla
- převrstvit převrstvovacím pufrem a dále postupovat jako v případě běžného gelu počínaje bodem 4

**Tabulka 10:** Použité objemy jednotlivých komponent pro přípravu roztoku dělicího gelu

Koncentrace gelu (%)	5	6	7	8	9	10	12	13	15	20
ddH <sub>2</sub> O	8,75	8,25	7,75	7,25	6,75	6,25	5,25	4,75	3,75	1,25
4x Tris-HCl/SDS pH 8,8	3,75	3,75	3,75	3,75	3,75	3,75	3,75	3,75	3,75	3,75
akrylamid + N,N' methylen- bisakrylamid	2,5	3	3,5	4	4,5	5	6	6,5	7,5	10
10 % (w/v) APS	0,075	0,075	0,075	0,075	0,075	0,075	0,075	0,075	0,075	0,075
TEMED	0,015	0,015	0,015	0,015	0,015	0,015	0,015	0,015	0,015	0,015

Uváděné objemy jsou v mililitrech. Objem dělicího roztoku připraveného podle této tabulky vystačí na přípravu dvou gelů.

#### 4.2.4.8.3 Elektroforetické dělení

- vytáhnout hřebínek, propláchnout jamky pomocí stříkačky naplněné 1x SDS-PAGE pufrem
- vložit skla s polyakrylamidovým gelem do elektroforetické aparatury, nalít dostatečné množství 1x SDS-PAGE pufru
- nanést vzorky smíchané se vzorkovým pufrem pro SDS-PAGE elektroforézu a inkubované v termobloku 5–10 min. při teplotě 100 °C
- zapojit elektroforetický přístroj do elektrického pole o konstantním napětí 175 V po dobu 1–1,5 hod.

#### 4.2.4.9 Western blot

- rozebrat elektroforetickou aparaturu, odstranit zaostřovací gel, dělicí gel inkubovat 15 min. v blotovacím roztoku na kývačce při pokojové teplotě
- na kazetu blotovacího přístroje Mini Trans-Blot<sup>®</sup> Electrophoretic Transfer Cell (BIO-RAD) položit vláknitou podložku (příslušenství aparatury), na ní naskládat 3 filtrační papíry Whatmann<sup>®</sup> 3MM, polyakrylamidový gel, nitrocelulózovou membránu o poréznosti 0,2 µm NC2 (Serva) nebo PVDF membránu\* o poréznosti 0,2 µm Immun-Blot<sup>®</sup> PVDF Membrane (BIO-RAD) a 3 filtrační papíry Whatmann<sup>®</sup> 3MM
- všechny vrstvy nejdříve namočit v blotovacím roztoku
- vytlačit bubliny zkumavkou a nvrch položit druhou vláknitou podložku, zavřít kazetu blotovacího přístroje a vložit do elektrodového modulu blotovacího přístroje
- elektrodový modul spolu s chladicí vložkou plnou ledu vložit do blotovací aparatury, nalít dostatečné množství blotovacího roztoku

- přístroj zapojit do elektrického pole o konstantním napětí 150 V po dobu 1–1,5 hod.
- rozebrat blotovací přístroj, tužkou označit pozici jednotlivých proteinů předbarveného standardu molekulových hmotností a stranu membrány, která byla v kontaktu s gelem
- gel obarvit pomocí *Coomassie Brilliant Blue G-250* viz protokol 4.2.4.11.1
- membránu inkubovat 1 hod. v blokovacím roztoku na kývačce při pokojové teplotě
- membránu inkubovat na otočném zrcadle přes noc při teplotě 4 °C zatavenou do fólie spolu s primární protilátkou naředěnou v blokovacím roztoku
- membránu promýt 2x 15 min. v odmyvacím roztoku na kývačce při pokojové teplotě
- membránu inkubovat 20 min. v blokovacím roztoku na kývačce při pokojové teplotě
- membránu inkubovat 2–3 hod. na otočném zrcadle při teplotě 4 °C zatavenou do fólie spolu se sekundární protilátkou naředěnou v blokovacím roztoku
- membránu promýt 3x 15 min. v odmyvacím roztoku na kývačce při pokojové teplotě
- analyzovat chemiluminiscenčně

\* při použití PVFD membrány nejprve PVDF membránu inkubovat 2 s v 100 % (v/v) methanolu při pokojové teplotě, opláchnout ddH<sub>2</sub>O a poté inkubovat 10–15 min. v blotovacím roztoku na kývačce při pokojové teplotě

#### 4.2.4.10 Chemiluminiscenční detekce

##### 4.2.4.10.1 Detekce pomocí přístroje ImageQuant™ LAS 4000

- připravit 2x roztok 4,5 ml ddH<sub>2</sub>O + 0,5 ml 1 M Tris-HCl, pH 8,5
- k jednomu roztoku přidat 3 µl 30 % (v/v) H<sub>2</sub>O<sub>2</sub>
- k druhému roztoku přidat 50 µl roztoku luminolu a 22 µl roztoku kyseliny p-kumarové
- těsně před použitím roztoky smíchat, roztokem pokrýt celou plochu membrány
- membránu založit do Eurofolie a na folii vyznačit fixem pozici jednotlivých proteinů předbarveného standardu molekulových hmotností
- membránu v Eurofolii vložit do přístroje ImageQuant™ LAS 4000 (GE Healthcare)
- nastavit chemiluminiscenční režim přístroje v dodaném programu
- zvolit vhodný režim citlivosti, rozlišení a času snímání vzhledem k intenzitě signálu, snímky uložit na pevný disk počítače
- nastavit *in vivo* režim přístroje v dodaném programu
- snímat pozici vyznačených proteinů předbarveného standardu molekulových hmotností, snímky uložit na pevný disk počítače

#### 4.2.4.11 Barvení polyakrylamidových gelů

##### 4.2.4.11.1 Barvení pomocí *Coomassie Brilliant Blue G-250*

- po rozebrání elektroforetického nebo blotovacího přístroje inkubovat gel 15–30 min. ve fixačním roztoku na kývačce při pokojové teplotě

- inkubovat gel přes noc v roztoku *Coomassie Blue* na kývačce při pokojové teplotě
- promývat dH<sub>2</sub>O na kývačce při pokojové teplotě do odbarvení gelu
- gel vložit do Eurofolie, naskenovat v rozlišení 300 bodů na palec a více, uložit na pevný disk počítače nebo pro získání vyšší citlivosti obarvit stříbrem viz protokol 4.2.4.11.2

#### 4.2.4.11.2 Rychlé barvení stříbrem

- gel obarvený pomocí roztoku *Coomassie Blue* inkubovat 10 min. na kývačce při pokojové teplotě v 50 ml 50 % (v/v) methanolu
- inkubovat 10 min. na kývačce při pokojové teplotě v 50 ml 5 % (v/v) methanolu
- inkubovat 10 min. na kývačce při pokojové teplotě v 50 ml roztoku 32 µM DTI
- inkubovat 10 min. na kývačce při pokojové teplotě v 50 ml 0,1 % (w/v) roztoku AgNO<sub>3</sub>
- inkubovat 1 min. na kývačce při pokojové teplotě v 50 ml ddH<sub>2</sub>O
- inkubovat v 50 ml vyvíjecího roztoku [3 % (w/v) Na<sub>2</sub>CO<sub>3</sub> + 1/2000 objemu 37 % (w/v) formaldehydu] dokud roztok nezačne hnědnout, poté okamžitě odstranit vyvíjecí roztok
- inkubovat v 50 ml vyvíjecího roztoku dokud roztok nezačne hnědnout, poté okamžitě odstranit vyvíjecí roztok
- inkubovat v 50 ml vyvíjecího roztoku téměř do požadovaného zbarvení proteinových pruhů
- okamžitě začít přisypávat bezvodou kyselinu citrónovou, dokud roztok s gelem šumí
- oplachovat 3x 15 min. na kývačce při pokojové teplotě v 50 ml ddH<sub>2</sub>O
- gel vložit do Eurofolie, naskenovat v rozlišení 300 bodů na palec a více, uložit na pevný disk počítače

### 4.2.5 Práce s viru podobnými částicemi

#### 4.2.5.1 Izolace viru podobných částic

##### 4.2.5.1.1 Izolace ultracentrifugací v gradientu chloridu cesného

Část tohoto protokolu vychází z izolace kvasinkových lineárních plasmidů ultracentrifugací do polštáře chloridu cesného podle (STAM *et al.* 1986).

- kvasinkovou kulturu nechat narůst tři dny při teplotě 28 °C v 200 ml média YPD
- centrifugovat 5 min./4000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat ve 100 ml ddH<sub>2</sub>O
- centrifugovat 5 min./4000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat ve 100 ml TESor pufru
- inkubovat 20 min. při pokojové teplotě
- centrifugovat 5 min./4000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat v 100 ml pufru TSor
- centrifugovat 5 min./4000g při pokojové teplotě

- odstranit supernatant, pelet resuspendovat v 30 ml pufru TSor + 10 mg lytikázy
- inkubovat 15 min. při pokojové teplotě
- centrifugovat 5 min./3000g při pokojové teplotě
- odstranit supernatant, pelet opatrně kývavě otáčivým pohybem resuspendovat v 50 ml pufru TSor
- centrifugovat 5 min./3000g při pokojové teplotě
- opakovat předchozí 2 kroky 2x
- dokonale odstranit veškerý supernatant, pelet zchladit na ledu 5 min.
- pelet opatrně kývavě otáčivým pohybem resuspendovat ve 13 ml ledového pufru TM + 1 tableta směsi inhibitorů proteáz cOmplete<sup>TM</sup> Mini (Roche) a 1 mM PMSF
- přenést pelet do Dounceova homogenizátoru vychlazeného na ledu, skleněným pístem pomalu pohybovat 3x nahoru a dolů
- centrifugovat 20 min./20000g při 4 °C
- supernatant odebrat do čistých mikrozkuhavek
- centrifugovat 20 min./25000g při 4 °C
- supernatant odebrat, doplnit do 12 ml pufrém TM a opatrně nanést na 1 ml 40 % (w/v) CsCl polštář v TM pufru v ultracentrifugační zkumavce (Beckman)
- centrifugovat 2 hod./35000 otáček za minutu při 10 °C v rotoru SW 41 Ti (Beckman) a ultracentrifuze Optima<sup>TM</sup> L-90K (Beckman)
- vzorky supernatantu, peletu a CsCl polštáře analyzovat pomocí agarosové a SDS-agarosové elektroforézy
- vzorek v CsCl polštáři doplnit do 10 ml 42 % (w/v) CsCl v TM pufru + 150 mM NaCl a přenést do ultracentrifugační zkumavky (Beckman)
- převrstvit cca 2 ml parařinového oleje bayol F
- centrifugovat 20 hod./35000 otáček za minutu při 18 °C v rotoru SW 41 Ti (Beckman) a ultracentrifuze Optima<sup>TM</sup> L-90K (Beckman)
- rozebrat gradient pomocí rozebírače frakcí Fraction Recovery System (Beckman) na 24 frakcí po cca 400 µl
- analyzovat pomocí agarosové a SDS-agarosové elektroforézy

#### 4.2.5.1.2 Izolace ultracentrifugací v gradientu sacharosy

- kvasinkovou kulturu nechat narůst tři dny při teplotě 28 °C v 200 ml média YPD
- centrifugovat 5 min./4000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat ve 100 ml ddH<sub>2</sub>O
- centrifugovat 5 min./4000g při pokojové teplotě
- odstranit supernatant, pelet resuspendovat ve 10 ml *GFP wash* + 5 mM MgCl<sub>2</sub>
- suspenzi rozdělit do čistých mikrozkuhavek po cca 800 µl, přidat 400–500 µl skleněných kuliček o průměru 0,45 mm
- inkubovat v oscilačním kulovém mlýnku MM301 (Retsch®) 5 min. při frekvenci 30 kmitů/s

- počkat, až se skleněné kuličky samovolně usadí na dně, odebrat supenzi buněk a pufru do čistých mikrozkušavek
- centrifugovat 5 min./8000g při teplotě 4 °C
- supernatant odebrat do čistých mikrozkušavek
- inkubovat 10 min. na ledu
- centrifugovat 20 min./20000g při teplotě 4 °C
- supernatant odebrat do čistých mikrozkušavek
- centrifugovat 20 min./25190g při 4 °C
- supernatant odebrat, doplnit do 9 ml pufrem *GFP wash* + 5 mM MgCl<sub>2</sub> a opatrně nanést na 4 ml 20 % (w/v) sacharosový polštář v *GFP wash* pufru + 5 mM MgCl<sub>2</sub> v ultracentrifugační zkumavce
- centrifugovat 2,5 hod./30000 otáček za minutu při 4 °C v rotoru SW 41 Ti (Beckman) a ultracentrifuze Optima™ L-90K (Beckman)
- vzorky supernatantu, peletu a sacharosového polštáře analyzovat pomocí agarosové a SDS-agarosové elektroforézy
- pelet dokonale resuspendovat v 800 µl *GFP wash* pufru + 5 mM MgCl<sub>2</sub>
- nanést na 12 ml lineárního gradientu 10–50% (w/v) sacharosy v *GFP wash* pufru + 5 mM MgCl<sub>2</sub> v ultraclear ultracentrifugační zkumavce, gradient připravit pomocí přístroje Gradient Master (Biocomp) a programu *long sucrose SW41 10–50, 11 steps*
- centrifugovat 3 hod./35000 otáček za minutu při 4 °C v rotoru SW 41 Ti (Beckman) a ultracentrifuze Optima™ L-90K (Beckman)
- rozebrat gradient pomocí UA-6 UV/Vis detector (ISCO, Inc.) na 24 frakcí po cca 500 µl vytlačováním gradientu 60% (w/v) sacharosou v *GFP wash* pufru rychlostí 2,2 ml/min. za stálého měření absorbance při 260 nm
- k frakcím přidat 10 µl (20 mg/ml) proteinázy K
- inkubovat 1,5 hod. při teplotě 50 °C
- přidat 500 µl upraveného fenolu, dokonale promíchat (zhruba polovina frakcí měla díky vysoké koncentraci sacharosy hustotu vyšší než fenol, a proto byly tyto vzorky naředěny a extrahovány ve větších objemech)
- centrifugovat 5 min./14000g při pokojové teplotě
- odebrat vrchní vodnou fázi, doplnit její objem do 500 µl TE pufrem a přidat k ní 250 µl upraveného fenolu a 250 µl upraveného chloroformu, dokonale promíchat
- centrifugovat 5 min./14000g při pokojové teplotě
- opakovat předchozí dva kroky 1x
- odebrat vrchní vodnou fázi, doplnit její objem do 500 µl TE pufrem a přidat k ní 500 µl upraveného chloroformu, dokonale promíchat
- centrifugovat 5 min./14000g při pokojové teplotě
- odebrat vrchní vodnou fázi, přidat k ní 0,1 objemu 3 M NaAc, 2,5 objemu 96 % (v/v) ethanolu a 1 µl (25 mg/ml) lineárního polyakrylamidu GenElute™
- inkubovat přes noc při –20 °C
- centrifugovat 10 min./14000g při pokojové teplotě
- odstranit supernatant, přidat 1 ml 70 % (v/v) ethanolu, dokonale promíchat
- centrifugovat 10 min./14000g při pokojové teplotě



- odstranit supernatant, pelet nechat dokonale vyschnout v termostatu při teplotě 37 °C
- rozpustit v 10 µl pufru TE
- analyzovat pomocí agarosové elektroforézy

#### 4.2.6 Práce s obrazovými daty

Digitalizovaná obrazová data gelů agarosových elektroforéz DNA či RNA, gelů polyakrylamidových elektroforéz radioaktivně značené DNA, gelů proteinových SDS-PAGE elektroforéz a snímků chemiluminiscenčně detekovaných proteinů metodou *western blot* zobrazená v této práci byla různě upravena. Úprava probíhala v programech Carestream (Kodak), Quantity One (BIO-RAD) a Adobe® Photoshop® CS5 (Adobe Systems Inc.). Vždy se jednalo vědecky akceptovatelné úpravy, které měly za cíl zpřehlednit zobrazovanou skutečnost, případně zobrazovanou skutečnost učinit lépe viditelnou. Mezi běžně použité úpravy patřilo oříznutí oblasti obrazu nezobrazující nic relevantního, otočení obrazu v prostoru, zrcadlové otočení obrazu v prostoru, odbarvení obrazu, invertování barev obrazu, změna velikosti obrazu, změna jasu a kontrastu obrazu. V programu Adobe® Photoshop® CS5 byly v některých případech upraveny křivky a úrovně obrazu.

#### 4.2.7 Bioinformatické analýzy

Pokud není uvedeno jinak, bylo použito výchozího nastavení příslušných programů.

##### **Příprava konstruktů DNA *in silico* a tvorba schémat vektorů:**

Clone Manager 9.2 Professional Edition (Scientific & Educational Software)

##### **Kontrola elektroforetogramu sekvenování DNA:**

ChromasPro 1.34 (Technelysium Pty Ltd)

##### **Prohledávání databází sekvencí DNA:**

BLASTN® (<http://blast.ncbi.nlm.nih.gov/>)

##### **Porovnávání sekvencí DNA:**

ClustalW 1.81 a manuálně v prostředí programu BioEdit 7.0.5.3 (HALL 1999)

##### **Tvorba sekvenčního loga DNA z porovnávaných sekvencí:**

WebLogo 2.8.2 (<http://weblogo.berkeley.edu/logo.cgi>) (CROOKS *et al.* 2004)

WebLogo 3.5.0 (<http://weblogo.threeplusone.com/create.cgi>)

##### **Statistická analýza porovnávaných sekvencí DNA:**

Seqool 3.1 (<http://www.biossc.de/seqool/download.html>) (WANG 2011)

### **Návrh sekvencí oligonukleotidových primerů:**

OligoAnalyser 3.1 (<http://eu.idtdna.com/analyzer/applications/oligoanalyzer/>)

Primer Blast (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>)

RealTime qPCR Assay Entry (<http://eu.idtdna.com/scitools/Applications/RealTimePCR/>)

### **Vyhodnocení real-time PCR:**

LightCycler<sup>®</sup> 480 Software, Verze 1.5 (Roche)

### **Predikce sekundární struktury RNA:**

RNAstructure (<http://rna.urmc.rochester.edu/RNAstructureWeb/>) (BELLAROUSOV *et al.* 2013)

### **Prohledávání databází sekvencí proteinů:**

BLASTP<sup>®</sup> (<http://blast.ncbi.nlm.nih.gov/>)

### **Porovnávání sekvencí proteinů:**

Clustal Ω (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) (SIEVERS *et al.* 2011)

manuálně v prostředí programu BioEdit 7.0.5.3 (HALL 1999)

### **Barevné zobrazení porovnání sekvencí proteinů:**

BoxShade 3.21 ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html))

### **Predikce terciární struktury proteinů homologním modelováním:**

Robetta (<http://rosetta.bakerlab.org/>) (KIM *et al.* 2004)

I-TASSER (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) (ZHANG 2008)

### **Prohledávání databáze terciárních struktur proteinů:**

DaliLite v. 3 ([http://ekhidna.biocenter.helsinki.fi/dali\\_server/start](http://ekhidna.biocenter.helsinki.fi/dali_server/start)) (HOLM a ROSENSTRÖM 2010)

### **Strukturní porovnávání terciárních struktur proteinů:**

DaliLite v. 3 ([http://ekhidna.biocenter.helsinki.fi/dali\\_server/start](http://ekhidna.biocenter.helsinki.fi/dali_server/start)) (HOLM a ROSENSTRÖM 2010)

The PyMOL Molecular Graphic System 1.3 (Schrödinger LLC.)

### **Vizualizace terciárních struktur proteinů:**

The PyMOL Molecular Graphic System 1.3 (Schrödinger LLC.)

### **Tvorba *maximum likelihood* fylogenetických stromů:**

PhyML 3.0 (<http://www.hiv.lanl.gov/content/sequence/PHYML/interface.html>) (GUINDON *et al.* 2010)

Použité nastavení: substituční model – LG, rovnovážné frekvence – odhadnout použitím substitučního modelu, proporce neměnných míst – odhadnout z dat, parametr gama tvaru – odhadnout z dat, počáteční strom – BioNJ, optimalizace počátečního stromu – topologie stromu a délka větví, vylepšení stromu – *subtree pruning a regrafting*, podpora větví – 1000 *bootstrap* opakování

### **Vizualizace fylogenetických stromů:**

Rainbow Tree

(<http://www.hiv.lanl.gov/content/sequence/RAINBOWTREE/rainbowtree.html>)

## 5 VÝSLEDKY

### 5.1 KVASINKOVÉ CYTOPLASMATICKÉ LINEÁRNÍ PLASMIDY

#### 5.1.1 SAM vazebné místo mRNA *capping* enzymu je nezbytné pro fungování pGKL plasmidů *in vivo*

Jak bylo uvedeno v kapitole 2.1.1.2, produkt genu *ORF3* plasmidu pGKL2 tvoří tripartitní mRNA *capping* enzym lineárních plasmidů. Protein K2ORF3p je esenciální pro udržení pGKL plasmidů v bunce a jeho RNA-trifosfátázová a RNA-guanylyltransferázová aktivita byla potvrzena *in vitro* (TIGGEMANN *et al.* 2001). Methyltransferázovou aktivitu proteinu K2ORF3p se však navzdory několika nezávislým pokusům nepodařilo *in vitro* prokázat (TIGGEMANN *et al.* 2001; VOPÁLENSKÝ 2007). To vedlo k otázce, zda je předpokládaná methyltransferázová doména K2ORF3p aktivní a potřebná pro udržení pGKL plasmidů v bunce.

##### 5.1.1.1 Příprava kmene s bodovými mutacemi methyltransferázové domény

Rozhodli jsme se testovat esencionalitu methyltransferázové (MT) domény K2ORF3p přípravou kmenu s rekombinantním plasmidem pGKL2 obsahujícím bodové mutace kodónů aminokyselinových zbytků K2ORF3p, nezbytných pro funkční methyltransferázovou aktivitu u již charakterizovaných mRNA *capping* enzymů. Pro alaninovou mutagenezi byly vybrány aminokyselinové zbytky D<sup>432</sup>, G<sup>438</sup> a D<sup>440</sup> K2ORF3p viz Obrázek 18. Stávající funkční metodika umožňující modifikaci kvasinkových cytoplasmatických genetických elementů využívá homologní rekombinaci *in vivo* (KÄMPER *et al.* 1989b). V této práci byly pro přípravu lineárních DNA fragmentů použitých k modifikaci pGKL plasmidů využity produkty PCR a fúzních PCR reakcí, obsahující na svých 5' a 3' koncích komplementární přesahy k oblastem na cíleném pGKL plasmidu. V případě fúzních PCR reakcí zpravidla produkt jedné dílčí PCR reakce kódoval gen pro rezistenci k antibiotiku umožňující následnou selekci a produkt druhé dílčí PCR reakce zpravidla kódoval modifikaci vybraného ORF.

Nejdříve byl namnožen pomocí první PCR reakce úsek 3' konce genu *K2ORF3* s primery ORF3\_SAMdel\_F1 a ORF3\_SAMdel\_R1, kterými byly zavedeny do výsledného produktu reakce mutace kódující D432A, G438A a D440A. Jako templát sloužily purifikované pGKL plasmidy. Pomocí druhé reakce PCR namnožen úsek odpovídající genu rezistence ke G418, před kterým se nacházela *K1ORF2* UCR [sekvence mezi iniciačním kodónem translace a UCS (včetně) příslušného ORF]. K reakci byly použity primery KL\_orf6C\_Flag2F

a ORF3\_SAMdel\_R2 a jako templát sloužil purifikovaný plasmid pRKL1-1<sup>1</sup>. Produkty obou reakcí s částečným vzájemným překryvem sloužily po purifikaci jako templát fúzní PCR reakce s primery ORF3\_SAMdel\_F1 a ORF3\_SAMdel\_R2. Produkt fúzní PCR reakce, tvořící kazetu integrující se do plasmidu pGKL2, byl izolován z agarosového gelu metodou 4.2.2.1.5 a vnesen do kmene *Kluyveromyces lactis* IFO1267 metodou 4.2.2.5.3. Transformované buňky byly vysety na misky YPDA s G418.

	motiv I	
Variola virus D1R	LAIDFENGAD	604
Moll. contagiosum virus MC090R	LAVDFGNGAD	709
Rabbit fibrioma virus D3R	LAIDFGNGAD	596
Mel. sanguinipes entomopoxvirus MSV067	LSIDIGRGGD	614
S. cerevisiae Abd1	LELGCGKGGD	178
Sch. pombe Pcm1	LDMGCGKGGD	153
C. albicans Abd1	LDLCCGKGGD	213
A. thaliana At3g20650	LDLACGKGGD	98
D. melanogaster l(2)35Bd	LDMCCGKGGD	149
X. laevis Rnmt	LDLGCCKGGD	137
M. musculus Rnmt	LDLGCCKGGD	200
E. cuniculi Ecm1	LDLGCCKGGD	78
K. lactis pGKL2 ORF3	LDIGSAKGGD	440
P. etchellsii pPE1B ORF3	LDIGSGKGGD	409
P. acaciae pPac1-1 ORF3	LDIGSGRGGD	405
S. kluyveri pSKL ORF3	LDIGSGKGGD	408
Vaccinia virus D1R esenciální zbytky	D G D	
H. sapiens Rnmt esenciální zbytky	D G D	
S. cerevisiae Abd1 esenciální zbytky	E G D	
E. cuniculi Ecm1 esenciální zbytky	D D	
Mutagenese K2ORF3	A A A	

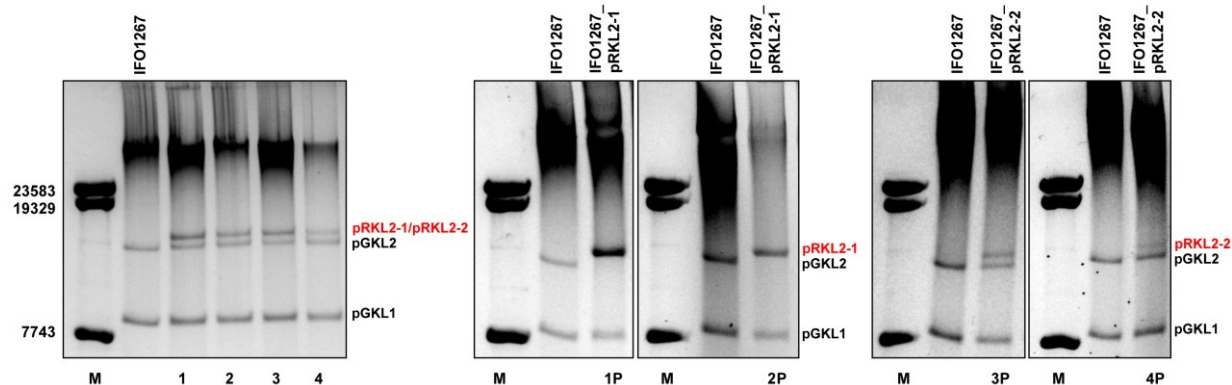
**Obrázek 18:** Sekvenční porovnání motivu I mRNA *cap* (guanin-N<sup>7</sup>)-methyltransferázových domén. Ve výsledném porovnání byla pomocí programu BoxShade 3.21 zvýrazněna sekvenční identita (černé stínování), pokud se stejný aminokyselinový zbytek vyskytoval u 60 a více procent sekvencí; sekvenční podobnost (šedé stínování), pokud se podobný aminokyselinový zbytek vyskytoval u 60 a více procent sekvencí. Pro tvorbu porovnání byly použity následující sekvence (číslo v závorce udává GenBank číslo dané sekvence): *Variola virus* D1R (P33057.1), *Molluscum contagiosum virus* MC090R (Q98257), *Rabbit fibrioma virus* D3R (P25950.1), *Melanoplus sanguinipes entomopoxvirus* MSV067 (Q9YW26), *Saccharomyces cerevisiae* Abd1 (P32783.1), *Schizosaccharomyces pombe* Pcm1 (O74880.2), *Candida albicans* Abd1 (Q5ADX5.1), *Arabidopsis thaliana* At3g20650 (Q9LHQ7.1), *Drosophila melanogaster* l(2)35Bd (Q9VJQ4.2), *Xenopus laevis* Rnmt (Q9I8S2.1), *Mus musculus* Rnmt (Q9D0L8.1), *Encephalitozoon cuniculi* Ecm1 (NP\_586153.1), *Kluyveromyces lactis* ORF3-pGKL2 (YP\_001648063.1), *Pichia etchellsii* ORF3-pPE1B (CAC08222.1), *Pichia acaciae* ORF3-pPac-1 (CAJ57276.1) a *Saccharomyces kluyveri* ORF3-pSKL (CAA38622.1). Pod porovnáním jsou zobrazeny esenciální aminokyselinové zbytky některých mRNA *cap* (guanin-N<sup>7</sup>)-methyltransferáz (MAO a SHUMAN 1996; WANG a SHUMAN 1997; SAHA *et al.* 1999; SCHWER *et al.* 2000; SAHA *et al.* 2003; HAUSMANN *et al.* 2005) a alaninová mutagenese K2ORF3p provedená v této práci. K alaninové mutagenesi K2ORF3p byly vybrány dva kyselé aminokyselinové zbytky, které jsou u ostatních mRNA *cap* (guanin-N<sup>7</sup>)-methyltransferáz velmi často esenciální, a také univerzálně konzervovaný glycinový zbytek.

<sup>1</sup> Rekombinantní plasmid *Kluyveromyces lactis* pGKL1 – 1 (druhé číslo slouží k odlišení jednotlivých rekombinantních pGKL plasmidů, připravených v naší laboratoři).

Několik vybraných monoklonií bylo otestováno na přítomnost rekombinantního plasmidu tak, že z nich byly pGKL plasmidy minipreparovány metodou 4.2.2.1.4 a podrobeny elektroforetickému dělení při napětí 1 V/cm minimálně po dobu 22 hodin. Buňky čtyř nezávislých klonů vykazujících přítomnost pGKL plasmidů divokého typu i modifikovaného plasmidu pGKL2 byly pasážovány v 5–10 ml tekutého média YPD + G418 s cílem vyředit plasmid pGKL2 divokého typu. Jedna pasáž trvala 12–24 hodin. Po páté (~60 generací) a desáté (~120 generací) pasáži byla část buněk vyseta na misky YPDA s G418 a u šestnácti monoklonií byla analyzována přítomnost pGKL plasmidů minipreparací a elektroforetickou analýzou. Modifikované plasmidy pGKL2 u čtyř výchozích nezávislých klonů a také několika odvozených klonů po pasážování byly izolovány pomocí metody 4.2.2.1.10 a jejich části byly amplifikovány pomocí PCR s primery in\_Kan\_rev1 + K2\_ORF3\_for\_seq. Produkty PCR byly rozděleny na agarosové elektroforéze, izolovány z gelu metodou 4.2.2.1.5 a podrobeny sekvenací primerem K2\_ORF3\_for\_seq.

U dvou zanalyzovaných výchozích klonů modifikovaný plasmid pGKL2 nazvaný pRKL2-1 neobsahoval požadované mutace v kódující sekvenci SAM vazebného motivu MT K2ORF3p. To bylo nejspíš způsobeno neúplnou rekombinací integrační kazety, protože úsek integrační kazety homologní ke genu *ORF3* byl poměrně dlouhý (~500 nt) a příslušné vnášené mutace se nacházely na úplném konci integrační kazety. U všech analyzovaných monoklonií obou klonů byl plasmid pGKL2 divokého typu vyředen již po 5. pasáži.

U dalších dvou zanalyzovaných výchozích klonů modifikovaný plasmid obsahoval požadované mutace v kódující sekvenci SAM vazebného motivu MT K2ORF3p. Tento plasmid byl nazván pRKL2-2. Z elektroforetogramu sekvenování však byla v menší míře patrná i přítomnost původních kodónů SAM motivu genu *K2ORF3*. Tyto klony tak byly pro vnášené mutace heterozygotní. To bylo nejspíš způsobeno současnou přítomností plasmidů pRKL2-1 a pRKL2-2 v těchto klonech. Jelikož se každý z pGKL plasmidů nachází v buňce přibližně v 50–100 kopiích (GUNGE 1983) může teoreticky dojít zároveň k úplné i neúplné integraci kazety v jedné buňce. Přibližně u poloviny analyzovaných monoklonií obou klonů po 5. a 10. pasáži, ve kterých byl plasmid pGKL2 divokého typu vyředen, byl sekvenací detekován výskyt plasmidu pRKL2-1. U druhé poloviny analyzovaných monoklonií obou klonů po 5. a 10. pasáži byl plasmid pGKL2 divokého typu zachován. U těchto klonů byl sekvenací detekován výskyt plasmidu pRKL2-2, přičemž některé z těchto klonů se podle detekčních limitů sekvenace jeví pro zaváděné mutace jako homozygotní. Z výše uvedeného je zřejmé, že plasmid pRKL2-2 kódující K2ORF3p s mutacemi SAM vazebného místa MT domény není schopen v buňce nahradit plasmid pGKL2 divokého typu. Elektroforetogramy detekce rekombinantních pGKL plasmidů v buňkách vybraných analyzovaných klonů před a po pasážování jsou na Obrázku 19.



**Obrázek 19:** Elektroforetogramy přípravy kmene IFO1267\_pRKL2-2, kódujícího K2ORF3p s mutacemi SAM vazebného místa MT domény, znázorňující detekci DNA pGKL plasmidů v buňkách analyzovaných klonů. Elektroforetogramy jsou zobrazeny v invertovaných barvách. Rekombinantní pGKL plasmidy jsou označeny červeně. M: Lambda DNA/Eco130I (*Sfi*I) Marker, 16; 1–4: nezávislé klony získané transformací kmene *K. lactis* IFO1267 divokého typu a následnou selekcí; 1P–4P: vybrané monoklonie získané po deseti pasážích nezávislých klonů 1–4. V případě monoklonií získaných pasáží klonu 3 a 4 obsahovala přibližně polovina kolonií plasmid pRKL2-2 a druhá polovina kolonií obsahovala plasmid pRKL2-1 a zároveň neobsahovala plasmid pGKL2 divokého typu. Velikost pGKL1 8874 bp, velikost pGKL2 13457 bp, velikost pRKL2-1 a pRKL2-2 14353 bp.

**Závěr:** SAM vazebné místo mRNA *capping* enzymu je nezbytné pro udržení pGKL plasmidů v buňce.

### 5.1.2 Komponenty transkripčního aparátu pGKL plasmidů interagují *in vivo*

Jak bylo uvedeno výše, geny *K2ORF3*, *K2ORF4*, *K2ORF6* a *K2ORF7* jsou nepostradatelné pro udržení pGKL plasmidů v buňce (SCHAFFRATH *et al.* 1995a, b; SCHAFFRATH *et al.* 1997; TIGGEMANN *et al.* 2001). Nicméně chápání interakcí proteinových produktů těchto genů mezi sebou a s plasmidovou DNA v buňce zcela chybí. Zpočátku jsme testovali interakce mezi K2ORF3p, K2ORF4p, K2ORF6p a K2ORF7p za použití kvasinkového dvouhybridního systému a jeho fluorescenční varianty nazvané bimolekulární komplementace fluorescence, ale těmito přístupy se nám nepodařilo prokázat žádnou interakci (VOPÁLENSKÝ 1999; SEKYROVÁ 2004; SÝKORA 2013). To bylo pravděpodobně způsobeno vysokým obsahem AT bází genů lineárních plasmidů, neboť nedávno bylo prokázáno, že jaderná exprese plasmidových genů není možná v důsledku fragmentace jejich mRNA zprostředkované buněčným štěpícím a polyadenylačním aparátem (KAST *et al.* 2015). Proto jsme se rozhodli připravit kmeny s modifikovanými pGKL plasmidy exprimujícími předpokládané komponenty transkripčního aparátu s různými proteinovými značkami, umožňujícími jejich purifikaci a detekci, pomocí výše popsané metodiky využívající homologní rekombinaci *in vivo*.



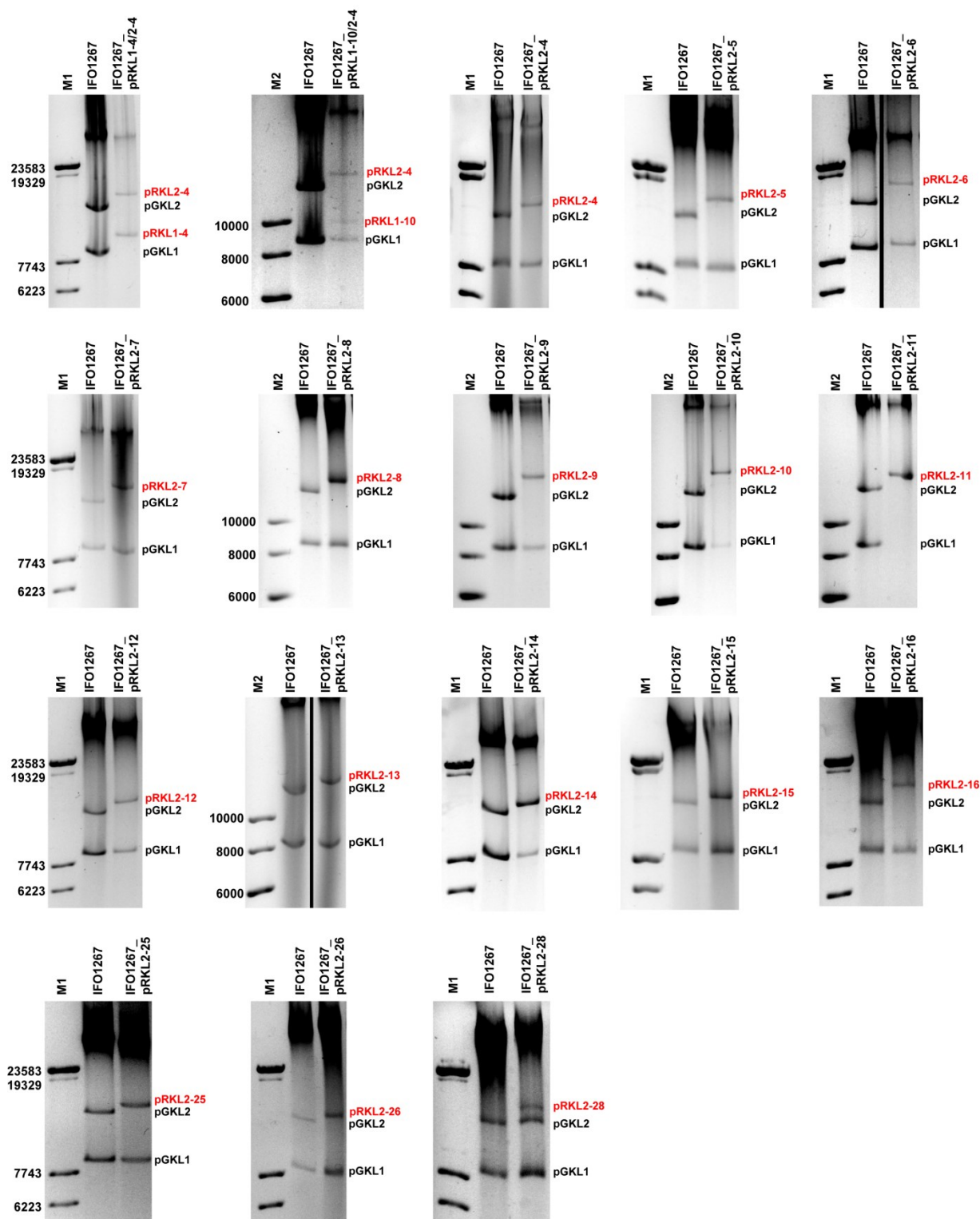
### 5.1.2.1 Příprava kmenů kódujících vybrané proteiny se značkami

Pro testování vzájemných interakcí proteinů transkripčního aparátu bylo připraveno velké množství kmenů (více než 18 kmenů) s jednoduše a dvojitě rekombinovanými pGKL plasmidy viz Tabulka 9. Příprava a charakterizace těchto kmenů tvořila časově nejobsáhlejší experimentální náplň této práce. Postup přípravy jednoduše rekombinovaných pGKL plasmidů je popsán v kapitole 5.1.1.1 na příkladu plasmidu pRKL2-2. Postup přípravy dvojitě rekombinovaných pGKL plasmidů spočíval v transformaci kmene *K. lactis* IFO1267 (metody 4.2.2.5.2 a 4.2.2.5.3) vytvořenou kazetou (metody 4.2.2.1.5, 4.2.2.1.6, 4.2.2.3.1, 4.2.2.6) kódující proteinovou značku ve fúzi s vybraným genem a rezistenci k jednomu antibiotiku. Získané analyzované rekombinantní klony (metody 4.2.2.1.4, 4.2.2.1.5, 4.2.2.1.10, 4.2.2.2, 4.2.2.3.1, 4.2.2.6, 4.2.2.8) byly pasážovány 60–120 generací v selektivním médiu (metoda 4.2.1.2). Pokud došlo po pasážování u analyzovaných klonů (metody 4.2.2.1.4, 4.2.2.1.5, 4.2.2.1.10, 4.2.2.2, 4.2.2.3.1, 4.2.2.6, 4.2.2.8) k vyředění cíleného plasmidu divokého typu, byly tyto nově získané kmeny transformovány kazetou kódující proteinovou značku ve fúzi s druhým vybraným genem a rezistenci k druhému antibiotiku. Získané klony byly pasážovány 60–120 generací v selektivním médiu. Pokud došlo po pasážování u analyzovaných klonů k vyředění jednoduše modifikovaného cíleného plasmidu, byly tyto klony použity k dalším experimentům (metody 4.2.4.4 – 4.2.4.11).

Celkový čas přípravy a charakterizace kmene s jednoduše a dvojitě rekombinovaným pGKL plasmidem je minimálně 1 a 2 měsíce. Elektroforetogramy kroků přípravy a charakterizace nových kmenů s rekombinantními pGKL plasmidy zde nejsou uvedeny. Z prostorových důvodů zde také nejsou uvedeny podrobné parametry PCR reakcí a fúzních PCR reakcí, použitých k přípravě transformačních kazet, stejně jako zde nejsou uvedeny parametry PCR reakcí sloužících k sekvenaci rekombinantních plasmidů a k detekci vyředění cíleného plasmidu divokého typu či jednoduše rekombinovaného cíleného plasmidu. Nejdůležitější informace k přípravě a charakterizaci nových kmenů souvisejících s touto kapitolou jsou shrnuty v Tabulce 11. Elektroforetogramy detekce DNA rekombinantních pGKL plasmidů v připravených kmenech jsou na Obrázku 20 a detailní popis charakteristik rekombinantních plasmidů je uveden v Tabulce 6. Všechny modifikované geny transkripčního aparátu, jejichž proteinové produkty byly studovány, byly ponechány pod kontrolou svých přirozených promotorů. Detekované proteinové interakce popsané v následující podkapitole tedy vycházejí z přirozených fyziologických koncentrací těchto proteinů v hostitelských buňkách.

**Tabulka 11: Primery a templáty pro přípravu a charakterizaci vybraných kmenů této práce**

Plasmid	Primery pro přípravu dílčích PCR reakcí	Templát pro přípravu dílčích PCR	Primery pro přípravu kazety pomocí fúzní PCR	Primery pro přípravu templátu na sekvenování	Primery pro ověření integrace kazety
				Sekvenace	Detekce vyředění nemodifik. plasmidu
pRKL1-4	K07 K1_ORF4-HA_R1	pGKL1	K07 K1_ORF4-HA_R2	Hygro_rev ORF4-pGKL1_tail	Hygro_rev 5RACE_O3_K1_3
	K1_ORF4-HA_F2 K1_ORF4-HA_R2	pcDNA™ 5/FRT/ TO		Hygro_rev	K08 ORF4-pGKL1_tail
pRKL1-10	5RACE_O3_K1_2 K1_HA-ORF3_R1	pRKL1-1	5RACE_O3_K1_2 K1_HA-ORF3_R2	nesekvenováno	K1_HA-ORF3_F2 5_RACE_O4_K1
	K1_HA-ORF3_F2 K1_HA-ORF3_R2	pcDNA™ 5/FRT/ TO		–	není vyředěn
pRKL2-4	KL_orf6N_Flag1F KL_orf6N_EGFP1R	pRKL1-1	KL_orf6N_Flag1F KL_orf6N_EGFP2R	5RACE_O6_K2 KanR1	5RACE_O6_K2 KanR1
	KL_orf6N_EGFP2F KL_orf6N_EGFP2R	pUG36		5RACE_O6_K2 KanR1	netestováno
pRKL2-5	KL_orf7C_Flag1F KL_orf7C_Flag1R	pGKL2	KL_orf7C_Flag1F KL_orf7C_Flag2R	ORF6-K2_tail_2 KL_orf7C_Flag2R	FLAG-rev ORF8-pGKL2_tail
pRKL2-15	KL_orf7C_Flag2F KL_orf7C_Flag2R	pcDNA™ 5/FRT/ TO		ORF7-pGKL2_tail	ORF6-K2_tail_2 ORF8-pGKL2_tail
pRKL2-6 pRKL2-10 pRKL2-14	pGKL_ORF3-HA_F1 pGKL_ORF3-HA_R1	pcDNA™ 5/FRT/ TO	pGKL_ORF3-HA_F1 K2_ORF3_for_seq	pGKL_ORF3-HA_F1 ORF3-K2_tail_3	pGKL_ORF3-HA_R1 5RACE_O2_K2
	pGKL_ORF3-HA_F2 K2_ORF3_for_seq	pGKL2		ORF3-K2_tail_3	5RACE_O2_K2 ORF3-K2_tail_3
pRKL2-7 pRKL2-13	pGKL_HA-ORF4_F1 pGKL_HA-ORF4_R1	pcDNA™ 5/FRT/ TO	pGKL_HA-ORF4_F1 5RACE_O4_K2	pGKL_HA-ORF4_F1 5RACE_O4_K2	pGKL_HA-ORF4_R1 ORF11_K2_tail_2
	pGKL_HA-ORF4_F2 5RACE_O4_K2	pGKL2		5RACE_O4_K2	5RACE_O4_K2 ORF11_K2_tail_2
pRKL2-8	EGFP-ORF4_1F kanR2	pRKL1-1	EGFP-ORF4_1F EGFP-ORF4_2R	KanVerF 5RACE_O4_K2	vORF3-k2-rev_2 in_Kan_rev1
	EGFP-ORF4_2F EGFP-ORF4_2R	pUG36		KanVerF 5RACE_O4_K2	vORF3-k2-rev_2 5RACE_O4_K2
pRKL2-9 pRKL2-12	pGKL_HA-ORF6_F1 pGKL_HA-ORF6_R1	pcDNA™ 5/FRT/ TO	pGKL_HA-ORF6_F1 in_ORF6_rev	pGKL_HA-ORF6_F1 5RACE_O6_K2	Hygro_rev in_ORF5_forw
	pGKL_HA-ORF6_F2 in_ORF6_rev	pGKL2		5RACE_O6_K2	k2orf6primex in_ORF5_forw
pRKL2-11	pGKL2_ORF3-EGFP_F1 KL_orf6C_Flag2F	pRKL1-1	pGKL2_ORF3- EGFP_F1 pGKL2_ORF3- EGFP_R2	in_Kan_rev1 ORF3-K2_tail_3	KanVerF 5RACE_O2_K2
	pGKL2_ORF3-EGFP_F2 pGKL2_ORF3-EGFP_R2	pUG36		in_Kan_rev1	netestováno
pRKL2-16	KL_orf7C_cMyc1F KL_orf7C_Flag1R	pGKL2	KL_orf7C_cMyc1F KL_orf7C_Flag2R	ORF6-K2_tail_2 KL_orf7C_Flag2R	c-myc_forw inORF6_forv2
	KL_orf7C_Flag2F KL_orf7C_Flag2R	pcDNA™ 5/FRT/ TO		ORF7-pGKL2_tail	ORF6-K2_tail_2 ORF8-pGKL2_tail
pRKL2-25	pGKL_HA-ORF6_F1 pGKL_HA-ORF6_R1	pcDNA™ 5/FRT/ TO	pGKL_HA-ORF6_F1 in_ORF6_rev	pGKL_HA-ORF6_F1 5RACE_O6_K2	pGKL_HA-ORF6_F1 5RACE_O6_K2
	pGKL_His-ORF6_F2 in_ORF6_rev	pGKL2		5RACE_O6_K2	k2orf6primex in_ORF5_forw
pRKL2-26	KL_orf6N_Flag1F pGKL2_GST-ORF6_R1	pRKL1-1	KL_orf6N_Flag1F 5RACE_O6_K2	KanVerF 5RACE_O6_K2	KanVerF 5RACE_O6_K2
	pGKL2_GST-ORF6_F2 5RACE_O6_K2	pGEX4T2/ GST-ORF6		KanVerF 5RACE_O6_K2	netestováno
pRKL2-28	ORF4-EGFP_1F ORF4-EGFP_1R	pUG36	ORF4-EGFP_1F ORF4-EGFP_2R	ORF4_K2_tail_2 in_Kan_rev1	KanVerF in_ORF5_rev
	KL_orf6C_Flag2F ORF4-EGFP_2R	pRKL2-1		ORF4_K2_tail_2 in_Kan_rev1	není vyředěn

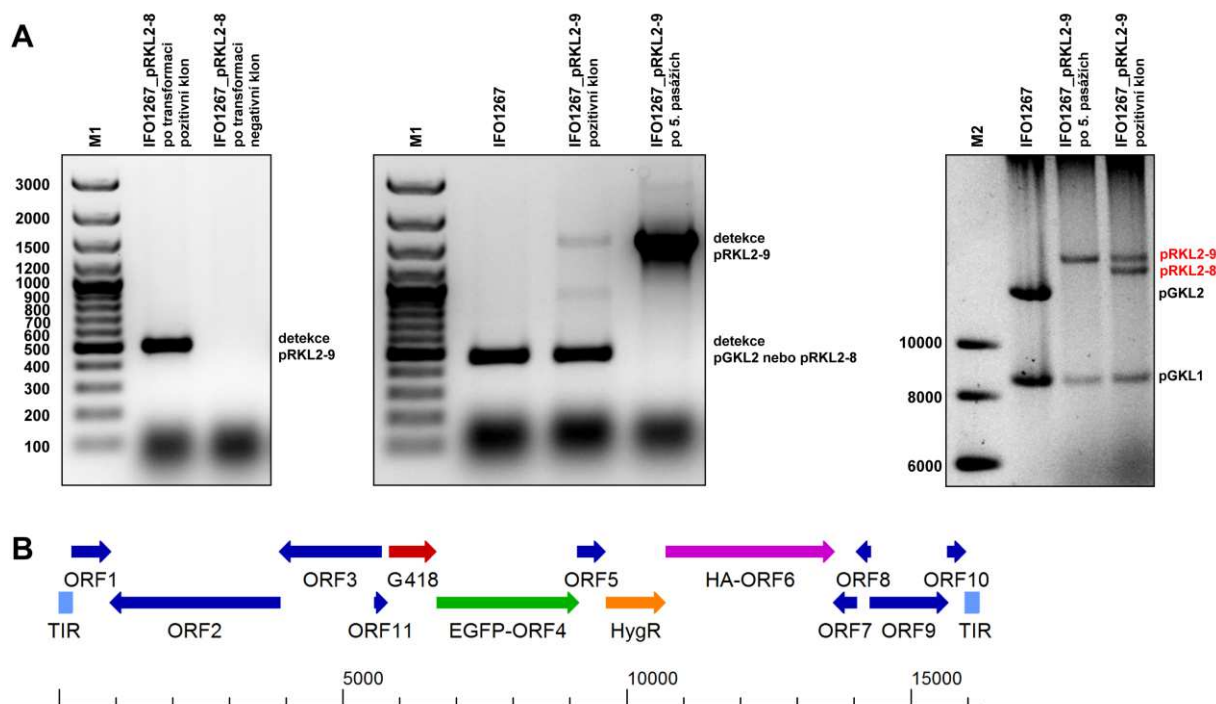


**Obrázek 20:** Elektroforetogramy připravených kmenů kódujících proteiny transkripčního aparátu pGKL plasmidů ve fúzi s různými proteinovými značkami. Rekombinantní pGKL plasmidy jsou označeny červeně. M1: Lambda DNA/Eco130I (*Sfi*I) Marker, 16; M2: GeneRuler™ 1 kb DNA Ladder. Zobrazeny jsou vybrané monoklonie po pěti či deseti pasážích nezávislých klonů získaných transformací a následnou selekcí. Kultivace kmene IFO1267\_pRKL2-11 v selektivních podmínkách vede u všech analyzovaných klonů k vyředění plasmidu pGKL1. Plasmid pRKL1-10 není schopen v buňkách zastoupit funkci plasmidu pGKL1 a plasmid pRKL2-28 není schopen v buňkách zastoupit funkci plasmidu pGKL2. Proto s kmeny IFO1267\_pRKL1-10/2-4 a IFO1267\_pRKL2-28 nebylo dále pracováno. Velikost rekombinantních pGKL plasmidů je uvedena v Tabulce 6.

Vě většině případů se po jistých optimalizacích podařilo pomocí dílčích PCR reakcí a následujících fúzních PCR reakcí připravit dostatečné množství integrační kazety. Pouze v případě přípravy plasmidů pRKL2-9, pRKL2-12 a pRKL2-25 vznikalo velmi malé množství finálního produktu fúzní PCR reakce. V těchto případech byl purifikovaný finální produkt reamplifikován pomocí PCR reakce využívající primery pro fúzní PCR reakci.

Charakterizace nových kmenů s rekombinantními pGKL plasmidy probíhala následujícím způsobem. Z klonů získaných transformací byly připraveny templáty pro PCR reakci, která by detekovala rekombinaci integrační kazety do cíleného pGKL plasmidu orientovaným způsobem. K tomu účelu byl vybrán jeden primer komplementární ke kazetě a druhý primer komplementární k plasmidu (Tabulka 11). U pozitivních klonů byly provedeny PCR reakce (Tabulka 11) pokrývající místo kódující zaváděnou proteinovou značku a tato oblast byla sekvenována pro ověření získané sekvence. V některých případech bylo testováno vyředění nemodifikovaného cíleného plasmidu i pomocí PCR reakcí. K tomu účelu byly vybrány dva primery komplementární jak k modifikovanému, tak k nemodifikovanému cílenému plasmidu (Tabulka 11), kdy v případě nemodifikovaného plasmidu vzniká kratší produkt než v případě modifikovaného plasmidu. U pozitivních klonů po transformaci byly agarosovou elektroforézou zpravidla detekovatelné oba PCR produkty či pouze kratší produkt, a po pasážování těchto klonů v případě vyředění nemodifikovaného plasmidu byl detekovaný majoritně či výhradně delší produkt. Touto analýzou bylo testováno, zda je v připravených kmenech naprostá většina studovaného proteinu označena zaváděnou značkou. Analýza vyředění nemodifikovaného cíleného plasmidu byla standardně prováděna minipreparací pGKL plasmidů a jejich elektroforetickou analýzou, nicméně dodatečná analýza pomocí PCR poskytovala větší citlivost. Pro ukázkou jsou uvedeny elektroforetogramy detekce rekombinace integrační kazety do cíleného plasmidu u klonů získaných po transformaci, detekce vyředění nemodifikovaného plasmidu na příkladu kmene IFO1267\_pRKL2-9 a schéma výsledného plasmidu pRKL2-9 (Obrázek 21).

Při přípravě kmenů bylo bráno v potaz, že na nativní C-konec proteinu K2ORF6p je esenciální pro udržení pGKL plasmidů v buňkách (SÝKORA 2013). Modifikace genů proteinů transkripčního aparátu pGKL plasmidů a vybraných kontrolních plasmidových proteinů většinou neměla vliv na soběstačnost vzniklých plasmidů, nicméně označení C-konce genu *K2ORF4* genem kódujícím EGFP a N-konce genu *K1ORF3* kódující sekvencí HA-značky vedlo k neschopnosti modifikovaných plasmidů nahradit v buňkách plasmidy divokého typu při inkubaci v selektivních podmínkách favorizujících modifikovaný plasmid viz Obrázek 20.



**Obrázek 21:** (A) Elektroforetogramy dokumentující přípravu a charakterizaci kmene *Khyveromyces lactis* IFO1267\_pRKL2-9, kódujícího gen *K2ORF4* ve fúzi se sekvencí kódující kvasinkový posílený zelený fluorescenční protein 3 (yEGFP3) (CORMACK *et al.* 1997) a gen *K2ORF6* ve fúzi se sekvencí kódující HA-tag. K detekci integrace kazety (první elektroforetogram zleva), nesoucí gen rezistence k hygromycinu B a HA-značku, byly použity primery Hygro\_rev a in\_ORF5\_forw (produkt 520 bp). K detekci vyředění nemodifikovaného cíleného plasmidu (druhý elektroforetogram zleva) byly použity primery k2orf6primex a in\_ORF5\_forw (produkty 489 bp a 1579 bp). Rekombinantní pGKL plasmidy jsou označeny červeně. M1: GeneRuler™ 100 bp Plus DNA Ladder; M2: GeneRuler™ 1 kb DNA Ladder. (B) Schématické znázornění kódující kapacity plasmidu pRKL2-9 (16194 bp). Šipky indikují otevřené čtecí rámce a směr jejich transkripce. Tyrkysové obdélníky indikují terminální invertované repetice (TIR). Zelená šipka indikuje otevřený čtecí rámec kódující rekombinantní protein K2ORF4p s posíleným zeleným fluorescenčním proteinem (EGFP) na N-konci a fialová šipka indikuje otevřený čtecí rámec kódující rekombinantní protein K2ORF6p se značkou HA na N-konci. Červená a oranžová šipka indikuje otevřený čtecí rámec kódující rezistenci ke G418 a hygromycinu B. U plasmidu je uvedeno měřítko v tisících párů bází.

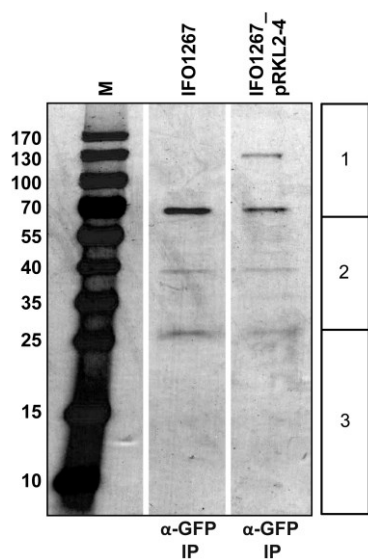
**Závěr:** Podařilo se zavést a optimalizovat systém umožňující přípravu a charakterizaci značného množství kmenů s modifikovanými pGKL plasmidy. N-konec genu *K1ORF3* a C-konec genu *K2ORF4* se nepodařilo označit *in vivo* způsobem, který by neinterferoval s biologickými funkcemi pGKL plasmidů, nutnými pro jejich udržení v hostitelských buňkách. Označení C-konce genu *K2ORF3* *in vivo* větší (EGFP), ne však malou (HA-tag), proteinovou značkou vede při inkubaci buněk v selektivních podmínkách favorizujících modifikovaný plasmid ke ztrátě plasmidu pGKL1.



### 5.1.2.2 Testování proteinových interakcí pomocí koimunoprecipitace a hmotnostní spektrometrie

Pro testování vzájemných interakcí složek transkripčního aparátu lineárních plasmidů kvasinek byly použity dva částečně odlišné přístupy. V prvním z nich bylo využito hmotnostní spektrometrie ve třech biologických replikátech.

V prvním experimentu byly z připraveného kmene kódující kvasinkový posílený zelený fluorescenční protein 3 (CORMACK *et al.* 1997) ve fúzi s N-koncem velké podjednotky RNAP pGKL plasmidů (yEGFP3-K2ORF6p; kmen *K. lactis* IFO1267\_pRKL2-4) a z kontrolního kmene, obsahujícího pGKL plasmidy divokého typu (IFO1267), připraveny proteinové lyzáty podle metody 4.2.4.4.2. Z proteinových lyzátů obou kmenů byl yEGFP3-K2ORF6p (dále jen EGFP-K2ORF6p) imunoprecipitován (metoda 4.2.4.6) za použití agarosových kuliček GFP-Trap®\_A, které obsahují lamí monoklonální protilátku proti běžným variantám GFP. Agarosové kuličky s navázanými proteiny byly třikrát promyty k odstranění slabě vázaných proteinů. Vázané proteiny byly rozděleny na SDS-PAGE proteinové elektroforéze (metoda 4.2.4.8) a obarveny pomocí *Coomassie Blue* (metoda 4.2.4.11.1). Úseky gelu odpovídající každému vzorku byly rozřezány na tři části (Obrázek 22) a samostatně podrobeny proteolytickému štěpení trypsinem a analýze pomocí hmotnostní spektrometrie (MS) na přístroji Orbitrap Velos Pro (Thermo Fisher Scientific) v Proteomic Core Facility v Evropské laboratoři molekulární biologie v Heidelbergu. Jednotlivé peptidy byly identifikovány Joannou Kirkpatrick za pomoci programu Mascot a nastavení minimálního mascot skóre peptidu 20 a hranice spolehlivosti identifikace proteinu 95 %. Celkem byly identifikovány peptidy 522 proteinů v případě kmene IFO1267 a 552 proteinů v případě kmene IFO1267\_pRKL2-4 (data neuvedena). Pouze v případě kmene IFO1267\_pRKL2-4 však byly detekovány peptidy odpovídající imunoprecipitované velké podjednotce RNAP EGFP-K2ORF6p a malé podjednotce RNAP K2ORF7p viz Tabulka 12.



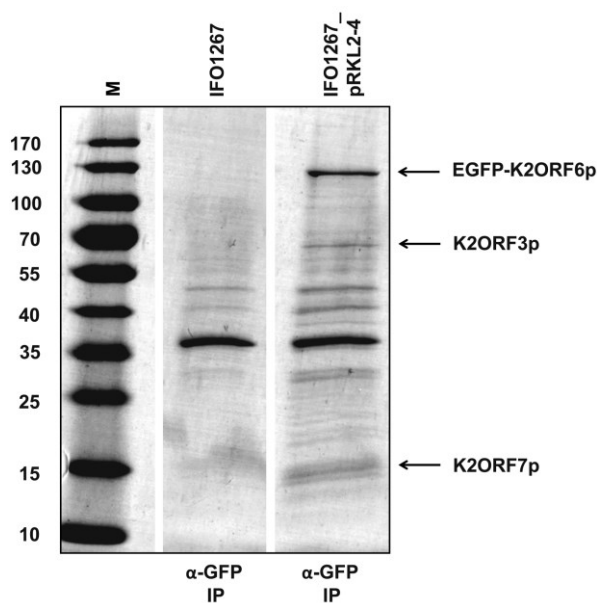
**Obrázek 22:** Identifikace proteinů asociovaných s velkou podjednotkou RNA polymerázy (K2ORF6p) lineárních plasmidů. Zobrazen je 5–20 % polyakrylamidový gel obarvený pomocí *Coomassie Blue*. Proteiny byly afinitně purifikované pomocí GFP-Trap®\_A (α-GFP IP) z kmenů IFO1267 (kontrola) a IFO1267\_pRKL2-4 (kódující EGFP-K2ORF6p). Vpravo jsou znázorněny a očíslovány úseky gelu, které byly rozřezány na části a samostatně analyzovány. Proteiny identifikované hmotnostní spektrometrií jsou uvedeny v Tabulce 12. Na gelu je patrný protein odpovídající molekulovou hmotností proteinu EGFP-K2ORF6p (140 kDa). M: PageRuler™ Prestained Protein Ladder.



Tabulka 12: Proteiny asociované s K2ORF6p				
Protein	kDa	Pokrytí	Počet peptidů	
			IFO1267_pRKL2-4	IFO1267
EGFP-K2ORF6p	140,7	48 %	170	0
K2ORF7p	15,5	53 %	22	0

Všechny další proteiny, které byly v tomto experimentu identifikované pouze u vzorku IFO1267\_pRKL2-4, měly poměrně malé množství detekovaných peptidů (průměrně 2,8 peptidu na protein), a proto nejsou považovány za opravdové interakční partnery velké podjednotky RNAP pGKL plasmidů. V tomto experimentu nebyly identifikovány další předpokládané komponenty transkripčního aparátu pGKL plasmidů, které by teoreticky mohly s RNA polymerázou interagovat. Při přípravě lyzátů kvasinkových protoplastů byla použita vysoká koncentrace chloridu draselného, která by mohla některé proteinové interakce narušit, proto bylo v dalším experimentu využito jiné přípravy proteinových lyzátů.

V druhém experimentu byly z kmene *K. lactis* IFO1267\_pRKL2-4 a z kontrolního kmene IFO1267 připraveny proteinové lyzáty podle metody 4.2.4.4.3. Z proteinových lyzátů obou kmenů byl EGFP-K2ORF6p imunoprecipitován pomocí GFP-Trap®\_A (metoda 4.2.4.6). Vázané proteiny byly rozděleny na SDS-PAGE proteinové elektroforéze (metoda 4.2.4.8) a obarveny pomocí *Coomassie Blue* (metoda 4.2.4.11.1). I v tomto případě byl na gelu u kmene IFO1267\_pRKL2-4 patrný protein odpovídající molekulovou hmotností proteinu EGFP-K2ORF6p, a navíc také proteiny odpovídající molekulovou hmotností proteinu K2ORF7p (15 kDa) a proteinu o přibližné molekulové hmotnosti 70 kDa (Obrázek 23). Protein odpovídající molekulovou hmotností proteinu EGFP-K2ORF6p a protein o přibližné molekulové hmotnosti 70 kDa byl vyřezán z gelu, odbarven, podroben štěpení trypsinem (SHEVCHENKO *et al.* 2006) a analýze pomocí hmotnostní spektrometrie na přístroji 4800 Plus MALDI TOF/TOF analyzer (AB Sciex) v Laboratoři hmotnostní spektrometrie Přírodovědecké fakulty Univerzity Karlovy. Jednotlivé peptidy byly identifikovány Karlem Harantem za pomoci programu Mascot 2.2.07. Paralelně byl protein o přibližné molekulové hmotnosti 70 kDa vyřezán z dalšího gelu, podroben štěpení trypsinem a analýze pomocí MS na přístroji Ultraflex III MALDI-TOF instrument (Bruker Daltonics) v Laboratoři charakterizace molekulární struktury Mikrobiologického ústavu Akademie věd ČR. Jednotlivé peptidy byly identifikovány Petrem Haladou za pomoci programu Mascot. Totožnost proteinu EGFP-K2ORF6p identifikovaného v předchozím experimentu byla potvrzena, a protein o přibližné molekulové hmotnosti 70 kDa byl v obou případech identifikován jako mRNA *capping* enzym pGKL plasmidů (K2ORF3p) viz Tabulka 13.



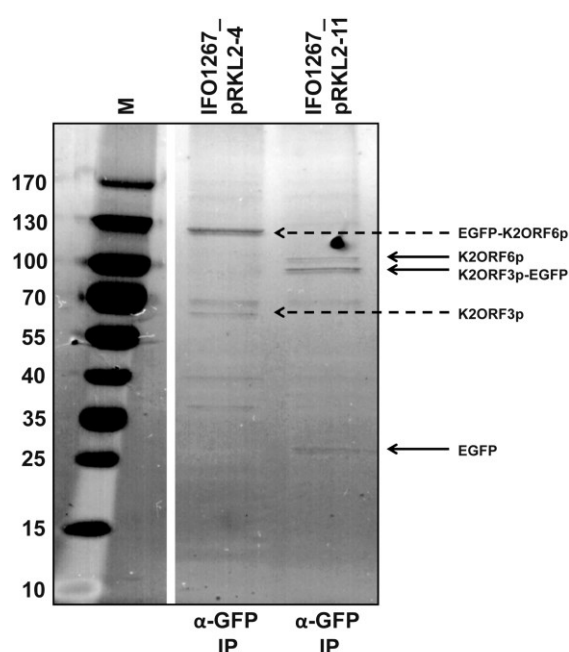
**Obrázek 23:** Identifikace proteinů asociovaných s velkou podjednotkou RNA polymerázy (K2ORF6p) lineárních plasmidů. Zobrazen je 5–20 % polyakrylamidový gel obarvený pomocí *Coomassie Blue*. Proteiny byly afinitně purifikované pomocí GFP-Trap®\_A ( $\alpha$ -GFP IP) z kmenů *K. lactis* IFO1267 (kontrola) a IFO1267\_pRKL2-4 (kódující EGFP-K2ORF6p). Vpravo jsou šipkou označeny proteiny vyřezané z gelu a identifikované hmotnostní spektrometrií v druhém experimentu. Počty detekovaných peptidů jsou uvedeny v Tabulce 13. Na gelu je patrný i protein odpovídající molekulovou hmotností proteinu K2ORF7p (15 kDa), jenž byl identifikován hmotnostní spektrometrií v předchozím experimentu viz Tabulka 12. M: PageRuler™ Prestained Protein Ladder.

**Tabulka 13:** Proteiny asociované s K2ORF6p

Protein	kDa	Pokrytí	Počet peptidů	Pracoviště
EGFP-K2ORF6p	140,7	17 %	33	PřF UK
K2ORF3p	70,5	24 %	13	PřF UK
K2ORF3p	70,5	23 %	11	AV ČR

Ve třetím experimentu byly z kmene *K. lactis* IFO1267\_pRKL2-4, IFO1267\_pRKL2-11 (K2ORF3p-yEGFP3) a z kontrolního kmene IFO1267 připraveny proteinové lyzáty podle metody 4.2.4.4.3. Z proteinových lyzátů tří kmenů byly imunoprecipitovány proteiny pomocí GFP-Trap®\_A (metoda 4.2.4.6). Vázané proteiny byly rovnou podrobeny štěpení trypsinem a analýze pomocí hmotnostní spektrometrie na přístroji Orbitrap Fusion™ Tribrid™ Mass Spectrometer (Thermo Fisher Scientific) v Proteomické servisní laboratoři centra BIOCEV. Vzorek z každého kmene byl na přístroji analyzován třikrát pro simulaci technického triplicátu. Jednotlivé peptidy byly identifikovány Karlem Harantem za pomoci programu Perseus 1.5.2.6 (COX a MANN 2012). Malá část vzorku vázaných proteinů byla rozdělena na SDS-PAGE proteinové elektroforéze (metoda 4.2.4.8) a obarvena stříbrem (metoda 4.2.4.11.2). V tomto případě byl na gelu u kmene IFO1267\_pRKL2-4 patrný protein odpovídající molekulovou hmotností proteinu EGFP-K2ORF6p a protein odpovídající molekulovou hmotností proteinu K2ORF3p (Obrázek 24). U kmene IFO1267\_pRKL2-11 byl na gelu patrný protein odpovídající proteinu K2ORF3p-EGFP a protein odpovídající proteinu K2ORF6p (Obrázek 24). Celkem byly identifikovány peptidy 455 proteinů v případě kmene IFO1267, 502 proteinů v případě kmene IFO1267\_pRKL2-4 a 490 proteinů v případě kmene IFO1267\_pRKL2-11 (data neuvedena). Většina proteinů (428) byla identifikována u všech tří vzorků. Proteiny pGKL

plasmidů detekované v jednotlivých vzorcích jsou uvedeny v Tabulce 14. V případě proteinů pGKL plasmidů, identifikovaných i v kontrolním kmeni IFO1267 (vztahuje se k proteinům K2ORF6p, K2ORF3p, K2ORF7p a jejich fúzním variantám), byl zpravidla počet detekovaných typů peptidů u IFO1267 menší než v modifikovaných kmenech, počet peptidů identifikovaných pomocí tandemové MS u IFO1267 menší než v modifikovaných kmenech, přičemž velká část detekovaných peptidů u IFO1267 nebyla detekována ve všech třech analýzách simulovaného technického triplikátu, na rozdíl od peptidů detekovaných u modifikovaných kmenů (data nezobrazena).



**Obrázek 24:** Identifikace proteinů asociovaných s mRNA *capping* enzymem lineárních plasmidů. Zobrazen je 5–20 % polyakrylamidový gel obarvený pomocí dusičnanu stříbrného. Proteiny byly afinitně purifikované pomocí GFP-Trap®\_A ( $\alpha$ -GFP IP) z kmenů *K. lactis* IFO1267\_pRKL2-4 (kódující EGFP-K2ORF6p) a IFO1267\_pRKL2-11 (kódující K2ORF3p-EGFP). Vpravo jsou přerušovanou šipkou označeny proteiny identifikované MS v druhém experimentu. Nepřerušovanou šipkou jsou označeny proteiny, jejichž molekulová hmotnost odpovídá proteinu K2ORF6p (113 kDa) a proteinu K2ORF3p-EGFP (98 kDa), které byly identifikovány pomocí MS ve třetím experimentu, viz Tabulka 14. Označen je i předpokládaný protein EGFP (29 kDa), jehož transkript může vznikat z promotoru K2UCR2, lokalizovaného na konci genu *K2ORF3*. Pruhy odpovídající malé podjednotce RNA polymerázy (K2ORF7p) nebyly jasně viditelné, pravděpodobně kvůli jejich menší hmotnosti a slabšímu barvení. M: PageRuler™ Prestained Protein Ladder.

Tabulka 14: Proteiny lineárních plasmidů asociované s K2ORF6p a K2ORF3p					
Protein	kDa	Pokrytí	Typy peptidů		
			IFO1267_pRKL2-11	IFO1267_pRKL2-4	IFO1267_pRKL2-11
EGFP-K2ORF6p	140,7	-	70	-	37
K2ORF6p	113,0	53 %	-	60	33
K2ORF3p-EGFP	98,2	43 %	-	37	22
K2ORF3p	70,5	-	27	-	18
K2ORF4p	66,9	2 %	1	1	0
K2ORF9p	55,0	11 %	5	6	0
K2ORF7p	15,5	74 %	10	9	6
K2ORF10p	12,1	10 %	2	2	2

U proteinů, jejichž peptidy byly identifikovány u vzorků všech tří kmenů (alespoň v jednom simulovaném technickém replikátu), byla provedena relativní kvantifikace vypočtením hodnoty MaxLFQ (COX *et al.* 2014) z peptidových signálů Karlem Harantem v programu Perseus 1.5.2.6 pro každý z těchto proteinů. Následně byla porovnána násobná změna (čtverec změny, *fold change*) rozdílu hodnot MaxLFQ mezi vzorky IFO1267\_pRKL2-4 a IFO1267, a IFO1267\_pRKL2-11 a IFO1267 pro tyto proteiny. Všechny proteiny, které vykazovaly *fold change* rozdílu hodnot MaxLFQ víc jak 20, a tedy největší pravděpodobné nabohacení ve vzorcích modifikovaných kmenů vůči kmeni IFO1267, jsou zobrazeny v Tabulce 15. U obou modifikovaných kmenů byl touto metodou správně identifikován jako nejvíc nabohacený ten protein, který byl imunoprecipitován (EGFP-K2ORF6p u kmene IFO1267\_pRKL2-4 a K2ORF3p-EGFP u kmene IFO\_pRKL2-11) viz Tabulka 15.

Tabulka 15: Proteiny asociované s K2ORF6p a K2ORF3p			
Protein	kDa	Čtverec změny ( <i>fold change</i> )	
		IFO1267_pRKL2-4 vs. IFO1267	IFO1267_pRKL2-11 vs. IFO1267
EGFP-K2ORF6p	140,7	139	-
K2ORF6p	113,0	-	35
K2ORF3p-EGFP	98,2	-	77
K2ORF3p	70,5	22	-
K2ORF7p	15,5	87	46
K2ORF10p	12,1	22	1

**Závěr:** Výsledky třech nezávislých experimentů imunoprecipitace a hmotnostní spektrometrie jasně ukazují, že podjednotky RNA polymerázy a mRNA *capping* enzym pGKL plasmidů asociují *in vivo* a pravděpodobně tvoří plasmidový transkripční komplex. Výsledky posledního experimentu ukazují, že s velkou podjednotkou RNAP pravděpodobně asociuje i plasmidový replikační protein TRF1 (K2ORF10p) *in vivo* a že s plasmidovým transkripčním komplexem možná *in vivo* asociuje i předpokládaná helikáza pGKL plasmidů (K2ORF4p) a plasmidový protein K2ORF9p se zatím neznámou funkcí. Výsledky také ukazují, že plasmidový transkripční komplex bude pravděpodobně soběstačný, protože se pomocí hmotnostní spektrometrie nepodařilo identifikovat žádný buněčný protein, který by se prokazatelně jevil specificky asociovaný s imunoprecipitovaným plasmidovým transkripčním komplexem.

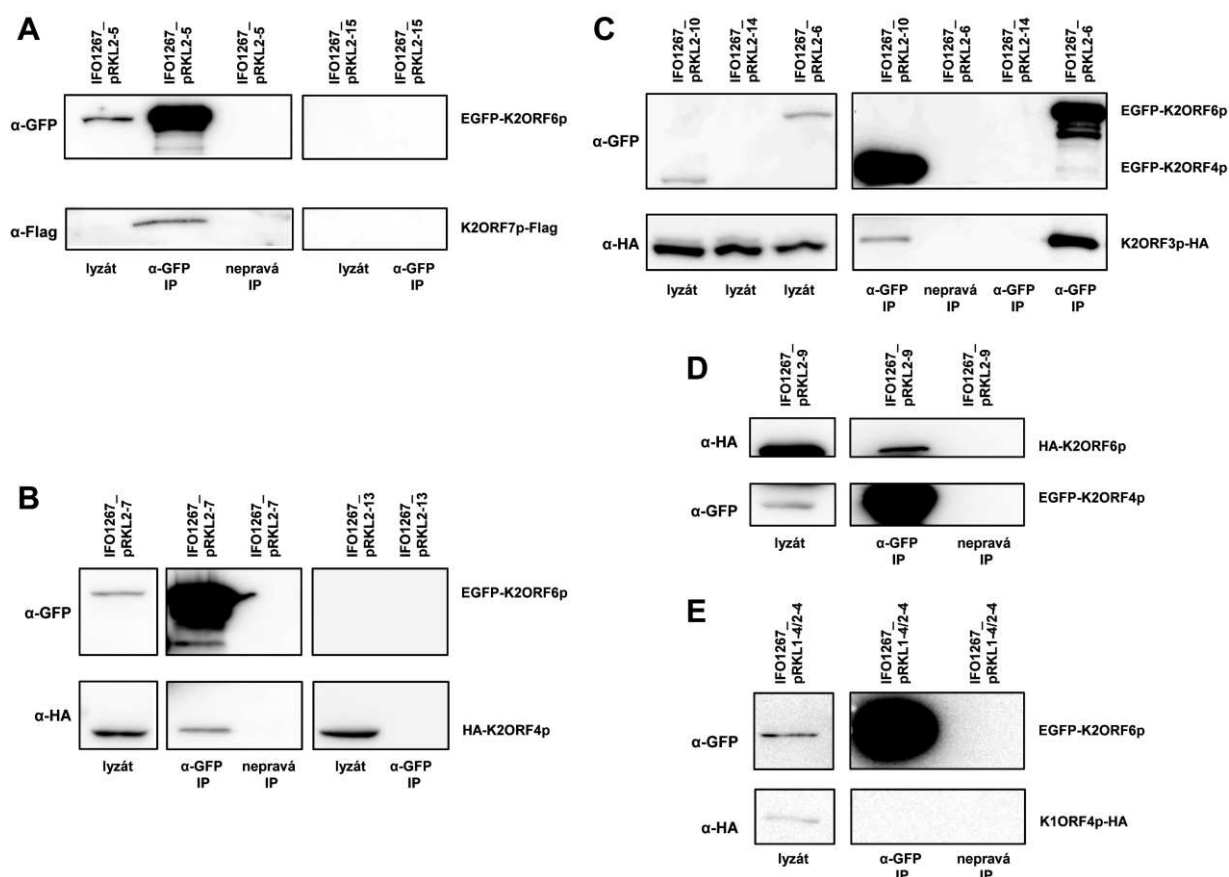
### 5.1.2.3 Testování proteinových interakcí pomocí koimunoprecipitace a imunodetekce

Pro validaci výsledků MS byla využita široká paleta připravených kmenů kódujících vždy dva vybrané proteiny transkripčního aparátu pGKL plasmidů s různými proteinovými značkami: (1) kmen *K. lactis* IFO1267\_pRKL2-5, kde byl EGFP-K2ORF6p exprimován společně s K2ORF7p-Flag, a kontrolní kmen IFO1267\_pRKL2-15, který expimoval pouze K2ORF7p-Flag; (2) kmen IFO1267\_pRKL2-6 exprimující současně EGFP-K2ORF6p a K2ORF3p-HA a kontrolní kmen IFO1267\_pRKL2-14 exprimující pouze K2ORF3p-HA. Z uvedených kmenů byly připraveny proteinové lyzáty (metoda 4.2.4.4.2) a z nich byl EGFP-K2ORF6p imunoprecipitován pomocí GFP-Trap®\_A nebo nepravě imunoprecipitován pomocí prázdných agarosových kuliček (metoda 4.2.4.6). Vzorky lyzátů a imunoprecipitovaných proteinů byly rozděleny na SDS-PAGE proteinové elektroforéze (metoda 4.2.4.8), přeneseny na membránu pomocí metody *western blot* (metoda 4.2.4.9), a vybrané proteiny byly chemiluminiscenčně imunodetekovány (metoda 4.2.4.10). Výsledky jasně ukázaly specifické interakce mezi podjednotkami RNAP a mRNA *capping* enzymem pGKL plasmidů *in vivo* (Obrázek 25A a C).

Předpokládaná helikáza K2ORF4p nebyla pomocí MS detekována jako součást transkripčního komplexu pGKL plasmidů. K ověření možné slabší interakce K2ORF4p s transkripčním komplexem byly připraveny kmeny exprimující HA-K2ORF4p společně s EGFP-K2ORF6p (IFO1267\_pRKL2-7) a kontrolní kmen pouze s HA-K2ORF4p (IFO1267\_pRKL2-13). Výsledky ukazují, že HA-K2ORF4p asociuje s EGFP-K2ORF6p (Obrázek 25C). Předpokládaná helikáza byla asociována s velkou podjednotkou RNAP specificky (nebo byla přítomna v komplexu obsahujícím tuto podjednotku RNAP), protože EGFP-K2ORF6p a HA-K2ORF4p nebyly vázány na prázdné agarózové kuličky a samotný protein HA-K2ORF4p se neváže na protilátku GFP-Trap®\_A (Obrázek 25C).

Vzhledem k tomu, že asociace předpokládané helikázy s velkou podjednotkou RNAP byla relativně slabá (signál v imunoprecipitovaném vzorku je slabší, než ve vzorku lyzátu) a pomocí MS nebyla příliš průkazná, byla provedena reciproká imunoprecipitace. Také byla testována možná interakce mezi mRNA *capping* enzymem a předpokládanou helikázou. Byly připraveny kmeny exprimující EGFP-K2ORF4p společně s HA-K2ORF6p (IFO1267\_pRKL2-9) a EGFP-K2ORF4p spolu s K2ORF3p-HA (IFO1267\_pRKL2-10). Při použití kmene IFO1267\_pRKL2-9 byla reciprokou imunoprecipitací potvrzena specifická interakce velké podjednotky RNAP s předpokládanou helikázou (Obrázek 25D). Při použití kmene IFO1267\_pRKL2-10 byla nalezena specifická interakce předpokládané helikázy s mRNA *capping* enzymem pGKL plasmidů (Obrázek 25B).





**Obrázek 25:** Interakce mezi podjednotkami RNA polymerázy, mRNA *capping* enzymem a předpokládanou helikázou pGKL plasmidů. (A) *Western blot* imunoprecipitovaných ( $\alpha$ -GFP IP) a nepravě imunoprecipitovaných (nepravá IP) proteinů z *K. lactis* IFO1267\_pRKL2-5 (EGFP-K2ORF6p, K2ORF7p-Flag) nebo kontrolních buněk IFO1267\_pRKL2-15 (K2ORF7p-Flag). Protilátky používané pro imunodetekci jsou uvedeny na levé straně výřezu. Detekované proteiny jsou uvedeny na pravé straně výřezu. Pozice identifikovaných proteinů odpovídaly teoretické molekulové hmotnosti rekombinantních proteinů plné délky, což bylo určeno podle pozic obarvených proteinů PageRuler™ Prestained Protein Ladder (nejsou zobrazeny). Lyzát představoval přibližně 1/100 vzorku a IP představoval přibližně 1/2 vzorku v tomto i v ostatních experimentech. Nepravá IP ve všech experimentech byla provedena použitím prázdných agarosových kuliček, pravá  $\alpha$ -GFP IP byla provedena pomocí GFP-Trap®\_A. Stejné experimentální schéma je použito i na dalších částech tohoto obrázku. (B) *Western blot* analýza imunoprecipitací a lyzátů IFO1267\_pRKL2-7 (EGFP-K2ORF6p, HA-K2ORF4p) a kontrolních buněk IFO1267\_pRKL2-13 (HA-K2ORF4p). (C) *Western blot* analýza imunoprecipitací a lyzátů z kmenů IFO1267\_pRKL2-6 (EGFP-K2ORF6p, K2ORF3p-HA), IFO1267\_pRKL2-10 (EGFP-K2ORF4p, K2ORF3p-HA) a IFO1267\_pRKL2-14 (kontrola). (D) *Western blot* analýza imunoprecipitací z lyzátů buněk IFO1267\_pRKL2-9 (EGFP-K2ORF4p, HA-K2ORF6p). (E) *Western blot* analýza imunoprecipitací a lyzátů buněk IFO1267\_pRKL1-4/2-4 (EGFP-K2ORF6p, K1ORF4p-HA). Pro detekci byly použity následující primární protilátky:  $\alpha$ -Flag M2 (ředění 1:1000),  $\alpha$ -HA 6E2 (1:1000) a  $\alpha$ -GFP B-2 (1:1000). Primární protilátky byly použity v kombinaci se sekundární protilátkou  $\alpha$ -mouse IgG-HRP (1:5000).

Pro další kontrolu, demonstrující specifitu pozorovaných interakcí, byl použit protein kódovaný pGKL plasmidy s biologickou funkcí přímo nesouvisející s transkripcí. Původně vybraný protein K1ORF3p kódující imunitní složku chránící hostitelské buňky pGKL plasmidů proti účinkům toxinu (TOKUNAGA *et al.* 1987) nebylo možné použít, neboť *in vivo* označení



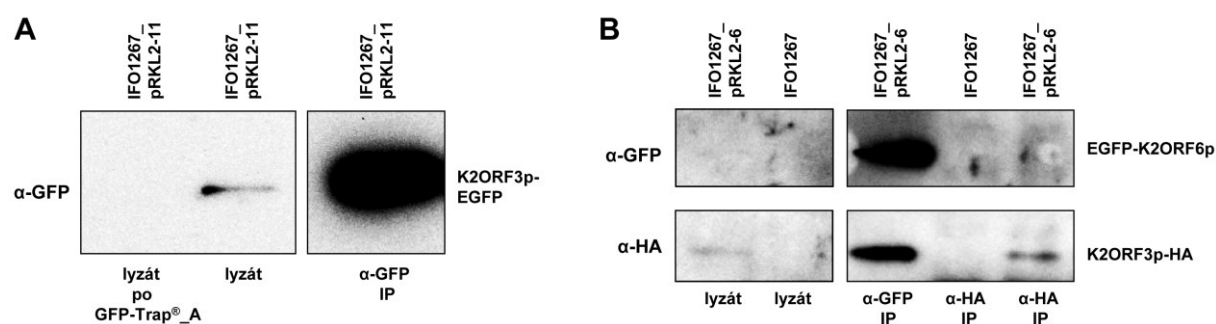
N-konce genu *K1ORF3* interferovalo s udržení rekombinantního plasmidu v buňkách (Obrázek 20). Následně byl vybrán K1ORF4p, což je  $\gamma$  podjednotka toxinu kódovaného pGKL plasmidy (TOKUNAGA *et al.* 1989). Byl připraven kmen koexprimující EGFP-K2ORF6p společně s K1ORF4p-HA (IFO1267\_pRKL1-4/2-4) a bylo potvrzeno, že  $\gamma$  podjednotka toxinu není asociována s velkou podjednotkou RNAP *in vivo* (Obrázek 25E).

Pro výše zmíněné experimenty se osvědčilo značení vybraných proteinů pomocí EGFP proteinu a HA-tagu, neboť v kombinaci s použitými protilátkami byly rekombinantní proteiny ve všech případech imunodetekovatelné jak ve výchozím lyzátu, tak v GFP-imunoprecipátu (Obrázek 25). Vybrané proteiny s HA-tagem bylo možné imunodetektovat pomocí protilátek anti-HA 6E2 a anti-HA HA-7, ale nebylo je možné imunodetektovat protilátkou anti-HA F-7 (data nezobrazena). Naproti tomu se značení vybraných proteinů pomocí Flag-tagu osvědčilo pouze částečně, protože v případě značené K2ORF7p-Flag byl rekombinantní protein pomocí protilátky anti-FLAG<sup>®</sup> M2 imunodetekovatelný pouze v GFP-imunoprecipátu a ne v lyzátu (Obrázek 25A). Obdobné výsledky byly dosaženy i v případě proteinu Flag-K2ORF6p (kmen IFO1267\_pRKL2-3), kdy rekombinantní protein není imunodetekovatelný v lyzátu a nedá se pomocí protilátky anti-FLAG<sup>®</sup> M2 imunoprecipitovat (SÝKORA 2013).

Neosvědčilo se značení proteinu K2ORF7p pomocí c-Myc-tagu (kmen IFO1267\_pRKL2-16), protože rekombinantní protein nebyl pomocí protilátky anti-c-Myc 9E10 imunodetekovatelný ani v lyzátu, ani v GFP-imunoprecipátu a nedal se pomocí protilátky anti-c-Myc 9E10 imunoprecipitovat (metoda 4.2.4.5, data nezobrazena). Pomocí protilátky anti-c-Myc 9E10 se však podařilo imunodetektovat kontrolní rekombinantní protein interleukin-1 $\alpha$ -c-Myc (data nezobrazena). Dále se také neosvědčilo značení K2ORF6p pomocí glutathion-S-transferázy (GST) a His-tagu (kmeny IFO1267\_pRKL2-25 a IFO1267\_pRKL2-26). Protein GST-K2ORF6p se nedal pomocí protilátky anti-GST 3D4 imunodetektovat v lyzátu, a nepodařilo se jej imunodetektovat ani při imunoprecipitaci, afinitní izolaci pomocí glutathion sepharosy a afinitní izolaci pomocí GST-Trap<sup>®</sup>\_A (metody 4.2.4.2, 4.2.4.5, 4.2.4.7, data nezobrazena). Protein His-K2ORF6p se nedal pomocí protilátky anti-polyHis HIS1 imunodetektovat v lyzátu a nepodařilo se jej detektovat při afinitní izolaci pomocí Ni-NTA agarosy (metoda 4.2.4.3, data nezobrazena).

Pro všechny výše zmíněné imunoprecipitační experimenty se vždy osvědčila imunoprecipitace EGFP-značeného proteinu pomocí GFP-Trap<sup>®</sup>\_A, který má vysokou vazebnou kapacitu pro GFP (2,5  $\mu$ g GFP/10  $\mu$ l GFP-Trap<sup>®</sup>\_A), a díky tomu funguje velice dobře a často i podstatně lépe než klasické imunoprecipitace. Ve výše zmíněných experimentech dokázal běžně GFP-Trap<sup>®</sup>\_A z lyzátu vyvázat všechny detekovatelný EGFP-značený protein (Obrázek 26A). Naproti tomu imunoprecipitace HA-značených proteinů pomocí anti-HA

agarosy HA-7 (Sigma) dosahovala podstatně menší účinnosti, jak dokládá Obrázek 26B, kde množství HA-imunoprecipitovaného proteinu K2ORF3p-HA je podstatně menší než množství GFP-koimunoprecipitovaného K2ORF3p-HA. V případě imunoprecipitace K2ORF3p-HA z kmene IFO1267\_pRK2-6 není v imunoprecipitátu patrný protein EGFP-K2ORF6p imunodetekcí (Obrázek 26B). Z uvedených důvodů byly reciproké imunoprecipitace vždy provedeny způsobem, že protein koprecipitující s EGFP-značeným proteinem byl v jiném kmeni sám označen pomocí EGFP a následně GFP-imunoprecipitován.



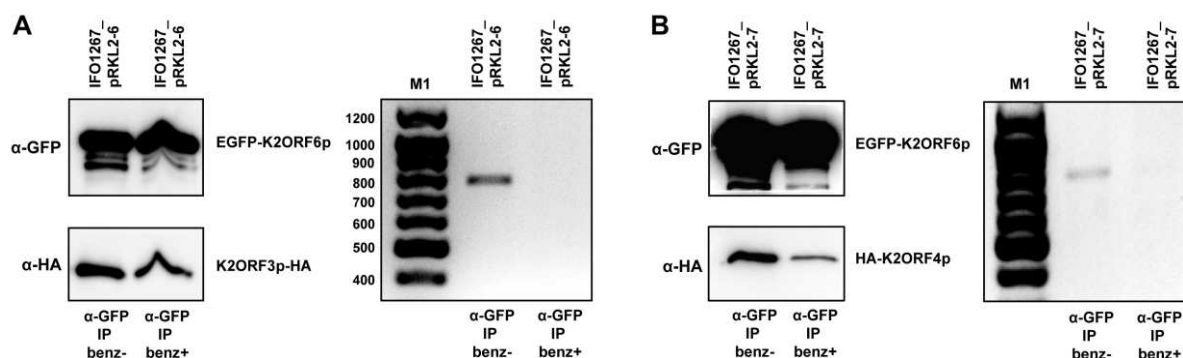
**Obrázek 26:** Vlastnosti GFP-Trap®\_A v použitém experimentálním uspořádání a porovnání s klasickou imunoprecipitací. (A) *Western blot* imunoprecipitovaných ( $\alpha$ -GFP IP) proteinů z *K. lactis* IFO1267\_pRK2-11 (K2ORF3p-EGFP). Protilátky používané pro imunodetekci jsou uvedeny na levé straně výřezu. Detekované proteiny jsou uvedeny na pravé straně výřezu. Pozice identifikovaných proteinů odpovídaly teoretické molekulové hmotnosti rekombinantních proteinů plné délky, což bylo určeno podle pozic obarvených proteinů PageRuler™ Prestained Protein Ladder (nejsou zobrazeny). Lyzát a lyzát po působení GFP-Trap®\_A představoval přibližně 1/100 vzorku a IP představoval přibližně 1/2 vzorku v tomto i v podobných experimentech. Imunoprecipitace  $\alpha$ -GFP byla provedena pomocí GFP-Trap®\_A. Stejné experimentální schéma je použito i na druhé části tohoto obrázku. (B) *Western blot* analýza lyzátů a reciprokých imunoprecipitací IFO1267\_pRK2-6 (EGFP-K2ORF6p, K2ORF3p-HA) a kontrolních buněk IFO1267. Imunoprecipitace  $\alpha$ -GFP byla provedena pomocí GFP-Trap®\_A, imunoprecipitace  $\alpha$ -HA byla provedena pomocí anti-HA agarosy HA-7. Pro detekci byly použity primární protilátky  $\alpha$ -HA 6E2 (ředění 1:1000) a  $\alpha$ -GFP B-2 (1:1000). Primární protilátky byly použity v kombinaci se sekundární protiláčkou  $\alpha$ -mouse IgG-HRP (1:5000).

**Závěr:** Podařilo se zavést a optimalizovat systém umožňující pomocí koimunoprecipitace a následné imunodetekce studovat interakce mezi vybranými proteiny kvasinkových lineárních plasmidů v jejich fyziologických koncentracích. Pomocí tohoto systému byla potvrzena těsná specifická interakce mezi podjednotkami RNA polymerázy a mRNA *capping* enzymu pGKL plasmidů *in vivo*, detekovaná již dříve pomocí hmotnostní spektrometrie. S předpokládaným transkripčním komplexem, poněkud slaběji, specificky interaguje i pravděpodobná DExD/H box helikáza pGKL plasmidů *in vivo*. Detekované interakce pravděpodobně reflektují biologickou roli zmíněných proteinů v procesu transkripce pGKL plasmidů, protože  $\gamma$  podjednotka toxinu s velkou podjednotkou RNAP *in vivo* dle očekávání neinteraguje.

### 5.1.2.4 Interakce transkripčních faktorů s RNA polymerázou není závislá na nukleových kyselinách

V dalším kroku jsme chtěli ověřit, zda je specifická interakce mRNA *capping* enzymu a pravděpodobné DExD/H box helikázy pGKL plasmidů s plasmidovou RNA polymerázou *in vivo* závislá na přítomnosti nukleových kyselin. Dříve detekované interakce složek transkripčního aparátu pGKL plasmidů nemusí nutně reflektovat protein–proteinové interakce, ale mohlo by se jednat o interakce zprostředkované přes nukleovou kyselinu, což by mohlo být možné vzhledem k předpokládaným RNA vazebným vlastnostem mRNA *capping* enzymu a DNA vazebným vlastnostem pravděpodobné helikázy.

Z proteinových lyzátů kmenů *K. lactis* IFO1267\_pRKL2-6 a IFO1267\_pRKL2-7 byl EGFP-K2ORF6p imunoprecipitován pomocí GFP-Trap®\_A (metoda 4.2.4.6). Navázané promyté proteiny byly ošetřeny benzonázou (nukleáza degragující DNA i RNA) v *GFP wash* pufru po dobu 30 minut při 33 ° C a znovu promyty před vlastní SDS-PAGE analýzou (metoda 4.2.4.8) a imunodetekcí pomocí metody *western blot* (metody 4.2.4.9 a 4.2.4.10). Část vázaných promytých proteinů byla odebrána a nukleové kyseliny byly eluovány povařením vzorku v TE pufru a detekovány amplifikací DNA pomocí PCR následované elektroforetickým dělením. Výsledky ukázaly, že interakce mezi velkou podjednotkou RNAP a mRNA *capping* enzymem či předpokládanou helikázou nejsou výrazně závislé na nukleových kyselinách (Obrázek 27).



**Obrázek 27:** Spojení předpokládané helikázy a mRNA *capping* enzymu s velkou podjednotkou RNAP pGKL plasmidů nezávisí na nukleových kyselinách. **(A)** EGFP-K2ORF6p byl purifikován pomocí GFP-Trap®\_A z lyzátů buněk *K. lactis* IFO1267\_pRKL2-6 (EGFP-K2ORF6p, K2ORF3p-HA). Po promytí navázaných proteinů (α-GFP IP) byly kuličky rozděleny na dvě části, které byly (benz+) nebo nebyly (benz-) ošetřeny 25 jednotkami benzonázy pro degradaci DNA a RNA. Po dalším promytí byly navázané proteiny eluovány a analyzovány pomocí *western blot*. Také byla odebrána část kuliček a navázané nukleové kyseliny byly eluovány a analyzovány na přítomnost DNA plasmidu pGKL1 PCR amplifikací (25 cyklů, primery 5\_RACE\_O2\_K1 a ORF1-K1\_tail\_3, produkt 785 bp) a následnou elektroforézou. Stejné experimentální schéma je použito i na druhé části tohoto obrázku. M1: GeneRuler™ 100 bp Plus DNA Ladder. **(B)** EGFP-K2ORF6p byl purifikován pomocí GFP-Trap®\_A z lyzátů buněk IFO1267\_pRKL2-7 (EGFP-K2ORF6p, HA-K2ORF4p). Část navázaných proteinů byla ošetřena 25 jednotkami benzonázy a navázané nukleové kyseliny byly eluovány a analyzovány na přítomnost DNA plasmidu pGKL1 PCR amplifikací (22 cyklů, stejné primery jako na Obrázku 27A). Pro imunodetekci byly použity primární a sekundární protilátky ve stejných ředěních, jako na Obrázku 26.

**Závěr:** Specifická interakce mRNA *capping* enzymu a pravděpodobné DExD/H box helikázy pGKL plasmidů s plasmidovou RNA polymerázou *in vivo* není závislá na nukleových kyselinách.

### 5.1.2.5 Předpokládné buněčné proteiny interagující s plasmidovým transkripčním komplexem

Jak již bylo uvedeno v kapitole 5.1.2.2, pomocí hmotnostní spektrometrie se nepodařilo identifikovat žádný buněčný protein, který by se prokazatelně jevil asociovaný s imunoprecipitovaným plasmidovým transkripčním komplexem. Ve třetím experimentu hmotnostní spektrometrie se však díky provedené relativní kvantifikaci detekovaných peptidů objevilo několik buněčných proteinů, které se zdály být ve vzorcích imunoprecipitátů modifikovaných kmenů oproti kmenu *K. lactis* IFO1267 nabohacené. Všechny buněčné proteiny, které vykazovaly *fold change* rozdílu hodnot MaxLFQ víc jak 3,5 u obou kmenů, či víc jak 5 u jednoho kmene jsou uvedeny v Tabulce 16.

Tabulka 16: Buněčné proteiny asociované s K2ORF6p a K2ORF3p				
Gen/Protein	Uniprot	kDa	Čtverec změny ( <i>fold change</i> )	
			IFO1267_ pRKL2-4 vs. IFO1267	IFO1267_ pRKL2-11 vs. IFO1267
eIF4G1	Q6CXY2	114,1	9,0	4,5
PABP1	Q6CSV3	66,0	4,1	3,7
RIM1	Q6CXR5	14,9	6,8	1,8
SBP1	Q6CQ18	27,4	3,6	3,5

Výsledky Tabulky 16 jsou poměrně překvapivé a mohly by být biologicky relevantní. Všechny čtyři proteiny jsou v databázi Uniprot anotovány jako ssDNA/RNA vazebné proteiny. Protein Rim1p *Saccharomyces cerevisiae*, což je předpokládaný homolog proteinu RIM1<sup>2</sup> (*Replication in mitochondria* 1) *K. lactis*, je jaderně kódovaný ssDNA vazebný protein, který je esenciální pro replikaci mitochondriální DNA *in vivo* (VAN DYCK *et al.* 1992). Buňky *Saccharomyces cerevisiae* (*S. cerevisiae*) s deletovaným genem *RIM1* vykazují fenotyp  $\rho^0$  (VAN DYCK *et al.* 1992). Tento protein byl nabohacený ve vzorku *K. lactis* IFO1267\_pRKL2-4, kde byl nalezen nabohacený plasmidový replikační protein TRF1. Je proto lákavá hypotéza, že by protein RIM1 mohl hrát roli v replikaci pGKL plasmidů.

Eukaryotický translační iniciační faktor 4 gama 1 (eIF4G1) a poly(A) vazebný protein 1 (PABP1) *S. cerevisiae* jsou ssRNA vazebné proteiny, které mají roli v translaci mRNA (SACHS

<sup>2</sup> Názvosloví především dosud necharakterizovaných genů/proteinů u *Kluyveromyces lactis* je poměrně nepřehledné, a proto jsem v tabulce použil funkční názvy předpokládaných paralogních genů/proteinů u *Saccharomyces cerevisiae*.

a DAVIS 1989; BURD *et al.* 1991; GOYER *et al.* 1993; BERSET *et al.* 2003). Oba tyto proteiny byly nabohaceny ve vzorcích z modifikovaných kmenů, a hypotézu o jejich možné asociaci s plasmidovým transkripčním aparátem, například zprostředkovaně přes RNA, umocňuje i fakt, že tyto dva proteiny v *S. cerevisiae* spolu interagují *in vivo* (TARUN a SACHS 1996). Sekvenční hledání možných paralogů proteinu KLLA0E00551p *K. lactis* (Uniprot číslo Q6CQ18) u *S. cerevisiae* pomocí BLASTP<sup>®</sup> odhalilo jako nejpravděpodobnější paralog protein Sbp1p (dříve známý jako *Single-stranded nucleic acid binding protein 1*, SSB-1). Protein Sbp1p *S. cerevisiae* má dvě RNA vazebné domény a ovlivňuje translační represi a odstraňování 5' mRNA čepiček (JONG a CAMPBELL 1986; SEGAL *et al.* 2006). Protein Sbp1p také interaguje s eIF4G1 (RAJYAGURU *et al.* 2012). Fakt, že tři nabohacené proteiny ve vzorcích imunoprecipitátů modifikovaných kmenů by teoreticky mohly interagovat *in vivo* v buňkách *K. lactis*, naznačuje možnost jejich interakce s plasmidovým transkripčním aparátem, hodnou dalšího výzkumu.

Pokusil jsem se ověřit předpokládanou interakci velké podjednotky RNAP pGKL plasmidů s translačním faktorem eIF4G1 koimunoprecipitací z kmene IFO1267\_pRKL2-4 (EGFP-K2ORF6p) pomocí GFP-Trap<sup>®</sup>\_A a následné imunodekce. Pomocí králíčího séra proti eIF4G1 *S. cerevisiae* se však nepodařilo imunodetektovat protein odpovídající svou molekulovou hmotností proteinu eIF4G1 u *K. lactis*, a to ani v lyzátu, ani v GFP-imunoprecipitátu, ačkoliv sérum rozeznávalo řadu buněčných proteinů v lyzátech modifikovaného i kontrolního kmene *K. lactis* IFO1267 (data nezobrazena).

**Závěr:** Výsledky hmotnostní spektrometrie ukazují, že s plasmidovým transkripčním komplexem mohou být asociovány i některé buněčné proteiny, ať už přímo či zprostředkovaně například přes nukleovou kyselinu. Pomocí koimunoprecipitace a následné imunodekce se zatím nepodařilo potvrdit či vyvrátit interakci velké podjednotky RNAP s eukaryotickým iniciačním translačním faktorem eIF4G1, což je podle výsledků kvantifikace peptidů detekovaných hmotnostní spektrometrií z buněčných proteinů nejvíce pravděpodobný interakční partner transkripčního komplexu pGKL plasmidů.

### 5.1.3 Produkce RNA polymerázy pGKL plasmidů v bakteriálním expresním systému není úspěšná

Je poměrně známo, že produkce proteinů kódovaných lineárními plasmidy kvasinek v klasických bakteriálních expresních systémech bývá problematická (SCHAFFRATH a MEACOCK 1995; TIGGEMANN *et al.* 2001). Díky velmi vysokému obsahu AT bází genů lineárních plasmidů



je používání kodónů jejich otevřených čtecích rámců dosti jiné od kodónů používaných v kódujících oblastech genů modelového organismu *Escherichia coli* (*E. coli*). Proto byl například protein K2ORF3p úspěšně produkován pouze v bakterii *Bacillus megaterium*, která má používáním kodónů k lineárním plasmidům blíže než bakterie *E. coli* (TIGGEMANN *et al.* 2001), a protein K2ORF5p byl po počátečních neúspěších s expresním systémem *E. coli* produkován v hmyzím bakulovirovém expresním systému (SCHAFFRATH a MEACOCK 1995). V naší laboratoři bylo uskutečněno již několik pokusů o produkci a purifikaci RNAP pGKL plasmidů v bakteriálních expresních systémech využívajících *E. coli*. V těchto systémech se podařilo produkovat His-K2ORF7p, i když purifikace většího množství čistého proteinu nebyla dokončena (VOPÁLENSKÝ 1999). Produkovat His-K2ORF6p ve stejném systému se podařilo jen ve velmi malém množství a protein se v buňkách rozpadá na několik degradačních produktů (VOPÁLENSKÝ 1999). V bakteriálním systému využívajícím buňky Rosetta, v kterých by měl být problém rozdílného používání kodónů částečně řešen produkcí několika transferových RNA, které jsou v *E. coli* raritní, se podařilo opakovaně produkovat GST-K2ORF7p, i když purifikace většího množství čistého proteinu nebyla dokončena (SÝKORA 2013). Ve stejném systému se však nepodařilo reprodukovat GST-K2ORF6p (SÝKORA 2013). Rozhodli jsme se otestovat, zda společná produkce obou podjednotek RNAP pGKL plasmidů v jednom bakteriálním kmeni z duálního expresního vektoru nepovede k lepší produkci, především rekombinantní velké podjednotky RNAP K2ORF6p.

### 5.1.3.1 Příprava duálních expresních vektorů

Pro společnou produkci podjednotek RNAP bylo třeba připravit vektory. Rozhodli jsme se vytvořit bicistronní vektory vložением genu *K2ORF7* či *K2ORF7-His* do vektoru pGEX4T2/TEV-ORF6 (kóduje gen GST-*K2ORF6* pod kontrolou Tac promotoru), a do vektoru pET16b/ORF6 (kóduje gen His-*K2ORF6* pod kontrolou T7/*lac* promotoru). Bicistronní transkripční kazeta byla navržena tak, že gen *K2ORF7* byl vložen za gen *K2ORF6* ve stejné transkripční orientaci před sekvenci terminátoru transkripce, a před jeho kódující sekvencí byla přítomna 5' nepřekládaná oblast odvozená z vektorů typu pET, obsahující translační enhancer genu *g10* bakteriofága T7 (OLINS a RANGWALA 1989) a Shine Dalgarno sekvenci AAGGAG, umístěnou 7 nukleotidů před iniciačním kodónem, což mělo umožnit translaci druhého cistronu.

Gen *K2ORF7* byl namnožen pomocí PCR s primery pGEX\_dual\_F a pGEX\_dual\_R či pGEX\_dual\_F a pGEX\_dual\_His\_R. Jako templát byl použit vektor pGEX4T2/GST-ORF7. Produkt reakce PCR byl izolován z gelu metodou 4.2.2.1.5. Přechištěný produkt byl štěpen

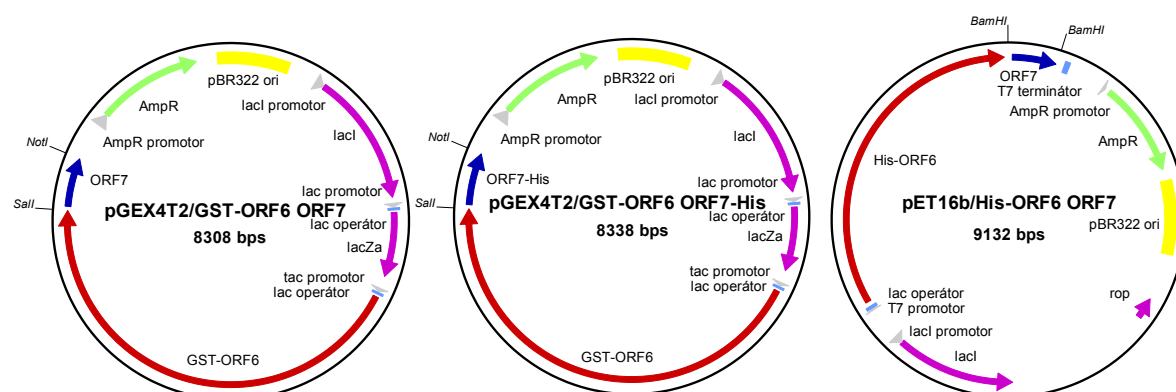


enzymy *Sall* a *NotI*, a posléze izolován z agarosového gelu (metoda 4.2.2.1.5). Naštěpený fragment byl ligován metodou 4.2.2.4.3 spolu s vektorem pGEX4T2/GST-ORF6 štěpeným za genem *K2ORF6* enzymy *Sall* a *NotI*, který byl po štěpení izolován z agarosového gelu metodou 4.2.2.1.5. Vektor pGEX4T2/GST-ORF6 jsem pro svoji práci nejdříve izoloval metodou 4.2.2.1.3. Ligační směs byla přečištěna (metoda 4.2.2.1.8) a část přečištěné ligační směsi byla vnesena do buněk *E. coli* XL-1 Blue elektroporací protokolem 4.2.2.5.1. Transformované buňky byly vysety na misky 2x TYA s ampicilinem. Vybrané monokolonie byly testovány na přítomnost rekombinantních vektorů tak, že z nich byla izolována DNA vektorů plasmidového typu minipreparací podle protokolu 4.2.2.1.1, a izolované vektory podrobeny restričnímu štěpení enzymy *Sall* a *NotI* s následnou elektroforetickou analýzou. Z buněk monokolonie, jejíž vektor po štěpení vykazoval přítomnost vektoru i insertu správné délky, byl vektor izolován metodou 4.2.2.1.2 a připraven k sekvenaci metodou 4.2.2.8.4 pomocí primerů 5RACE\_O7\_K2 a ORF7-K2\_tail\_2. Sekvence odhalila správné napojení insertu a požadované nukleotidové pořadí celého genu *K2ORF7*. Takto se podařilo připravit vektory pGEX4T2/GST-ORF6 ORF7 a pGEX4T2/GST-ORF6 ORF7-His.

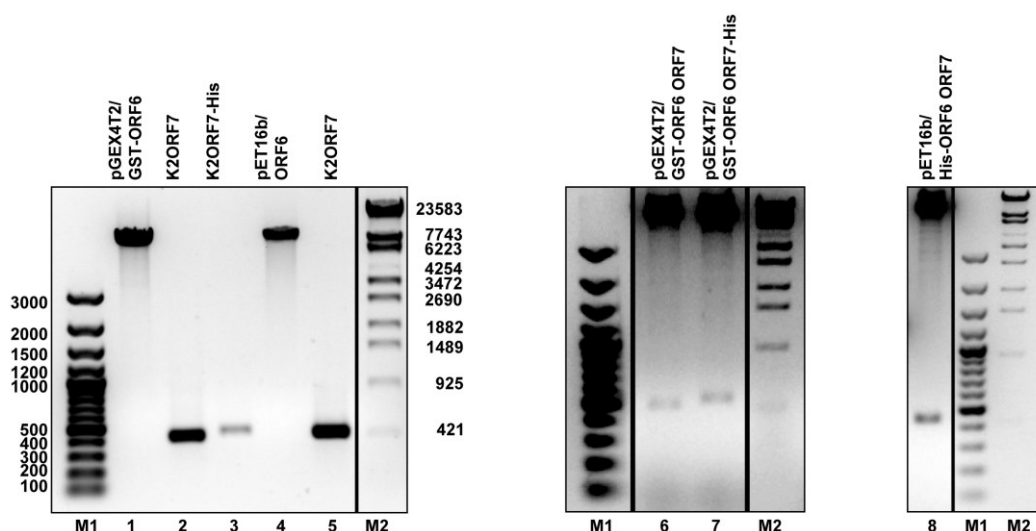
Namnožený gen *K2ORF7* výše uvedenými primery byl po izolaci také štěpen enzymem *BamHI*, a posléze izolován z agarosového gelu (metoda 4.2.2.1.5). Naštěpený fragment byl ligován (metoda 4.2.2.4.3) spolu s vektorem pET16b/ORF6 štěpeným za genem *K2ORF6* enzymem *BamHI*, který byl po štěpení izolován z agarosového gelu metodou 4.2.2.1.5. Ligační směs byla přečištěna (metoda 4.2.2.1.8) a část přečištěné ligační směsi byla vnesena do buněk *E. coli* XL-1 Blue elektroporací protokolem 4.2.2.5.1. Selektované monokolonie byly testovány na přítomnost rekombinantních vektorů tak, že z nich byla izolována DNA vektorů plasmidového typu minipreparací podle protokolu 4.2.2.1.1, a izolované vektory podrobeny restričnímu štěpení enzymy *BamHI* s následnou elektroforetickou analýzou. Z buněk monokolonie, jejíž vektor po štěpení vykazoval přítomnost vektoru i insertu správné délky, byl vektor izolován metodou 4.2.2.1.2 a připraven k sekvenaci metodou 4.2.2.8.4 pomocí primeru 5RACE\_O7\_K2. Sekvence odhalila správnou orientaci a napojení insertu a požadované nukleotidové pořadí dvou třetin genu *K2ORF7*. Takto se podařilo připravit vektor pET16b/His-ORF6 ORF7. Vektor pET16b/His-ORF6 ORF7-His se během tří opakovaných pokusů o ligaci připravit nepodařilo.

Připravené vektory pGEX4T2/GST-ORF6 ORF7 a pGEX4T2/GST-ORF6 ORF7-His izolované metodou 4.2.2.1.2 byly vneseny do buněk *E. coli* Rosetta<sup>TM</sup>(DE3) elektroporací. Připravený vektor pET16b/His-ORF6 ORF7 izolovaný metodou 4.2.2.1.2 byl vnesen do buněk *E. coli* Rosetta 2(DE3)pLysS elektroporací. Transformované buňky byly vysety na misky 2x TYA s ampicilinem a chloramfenikolem. V narostlých buňkách tří transformovaných kmenů byla

indukována exprese rekombinantních proteinů a pokusil jsem se o detekci a purifikaci těchto proteinů viz následující kapitola. Schéma vektorů je zobrazeno na Obrázku 28. Elektroforetogramy některých kroků přípravy vektorů jsou zobrazeny na Obrázku 29.



**Obrázek 28:** Schéma připravených duálních expresních vektorů pro produkci obou podjednotek RNA polymerázy pGKL plasmidů v bakteriálním expresním systému. Gen *K2ORF7* a *K2ORF7-His* byl namnožen pomocí PCR a štěpen restrikčními endonukleázami *SalI* a *NotI* či *BamHI*, a vložen do stejné štěpeného vektoru pGEX4T2/GST-ORF6 či pET16b/ORF6. V připravených vektorech je exprese rekombinantních proteinů z bicistronní kazety pod kontrolou Tac či T7 promotoru a *lac* operátoru. U každého vektoru je uvedena jeho velikost v párech bází. AmpR: gen rezistence k ampicilinu; lacI: lac represor; lacZa:  $\alpha$ -peptid  $\beta$ -galaktosidasy.



**Obrázek 29:** Elektroforetogramy některých kroků přípravy duálních expresních vektorů pro produkci RNAP pGKL plasmidů v bakteriálním expresním systému. U standardů molekulových hmotností jsou vyznačeny velikosti některých pruhů v párech bází. M1: GeneRuler™ 100 bp Plus DNA Ladder; M2: Lambda DNA/Eco130I (*StyI*) Marker; 16; 1: Vyizolovaný fragment vektoru pGEX4T2/GST-ORF6 štěpeného enzymy *SalI* a *NotI*, velikost 7876 párů bází (bp); 2 a 3: Vyizolovaný fragment genu *K2ORF7* a *K2ORF7-His* namnožený pomocí PCR a štěpený enzymy *SalI* a *NotI*, velikost 432 a 462 bp; 4: Vyizolovaný fragment vektoru pET16b/ORF6 štěpeného enzymem *BamHI*, velikost 8687 bp; 5: Vyizolovaný fragment genu *K2ORF7* namnožený pomocí PCR a štěpený enzymem *BamHI*, velikost 445 bp; 6: Vektor pGEX4T2/GST-ORF6 ORF7 štěpený enzymy *SalI* a *NotI*; 7: Vektor pGEX4T2/GST-ORF6 ORF7-His štěpený enzymy *SalI* a *NotI*; 8: Vektor pET16b/His-ORF6 ORF7 štěpený enzymem *BamHI*.

**Závěr:** Většinu plánovaných duálních vektorů pro expresi RNA polymerázy pGKL plasmidů v bakteriálním systému se podařilo vytvořit.

### 5.1.3.2 Purifikace a detekce rekombinantních proteinů

Produkce rekombinantních proteinů ze tří duálních vektorů v expresních buňkách *E. coli* byla vyzkoušena jednou na 100 ml vstupní kultury. V buňkách byla indukována exprese podle protokolu 4.2.1.5, a potom byly z buněk připraveny proteinové lyzáty podle protokolu 4.2.4.1. Z lyzátů byl protein GST-K2ORF6p purifikován metodou 4.2.4.2 a His-K2ORF6p byl purifikován metodou 4.2.4.3. Vzorky lyzátů a některých kroků purifikace byly rozděleny SDS-PAGE proteinovou elektroforézou (metoda 4.2.4.8), proteiny byly přeneseny na membránu pomocí metody *western blot* (metoda 4.2.4.9) a rekombinantní proteiny byly imunodetekovány (metoda 4.2.4.10) s použitím primárních protilátek anti-GST 3D4 a anti-polyHis HIS1 (ředění 1:1000) v kombinaci se sekundární protilátkou anti-mouse IgG-HRP (ředění 1:5000).

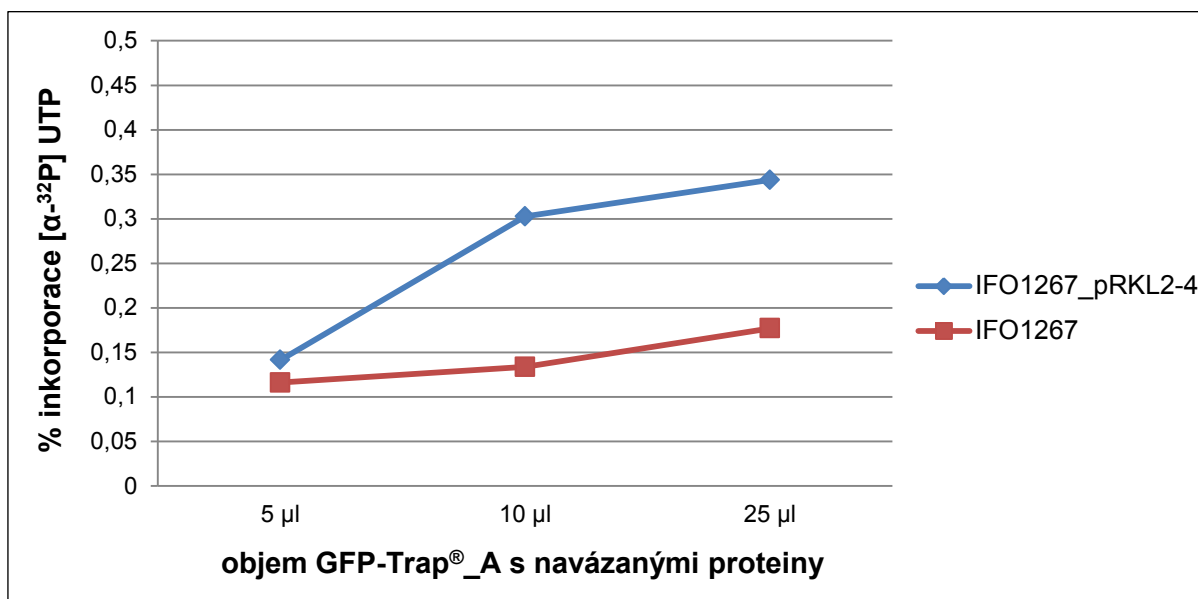
V lyzátech expresních buněk se podařilo detekovat velké množství proteinů různých molekulových hmotností, ale rekombinantní proteiny, které by svojí molekulovou hmotností mohly odpovídat rekombinantním proteinům GST-K2ORF6p, His-K2ORF6p či K2ORF7p-His, se nepodařilo spolehlivě identifikovat (data nezobrazena). Pomocí purifikace GST-fúzních proteinů se nepodařilo purifikovat a detekovat protein GST-K2ORF6p plné délky, ačkoliv se podařilo purifikovat několik proteinů menší než očekávané molekulové hmotnosti, z nichž některé by mohly odpovídat degradačním produktům proteinu GST-K2ORF6p (data nezobrazena). Pomocí purifikace GST-fúzních proteinů a imunodetekce se nepodařilo kopurifikovat a detekovat protein K2ORF7p-His (data nezobrazena). Pomocí purifikace proteinů s His-tagem se nepodařilo purifikovat a detekovat protein His-K2ORF6p plné délky, ačkoliv se podařilo purifikovat protein o přibližné molekulové hmotnosti 40 kDa, který by mohl odpovídat degradačnímu produktu rekombinantního proteinu His-K2ORF6p (data nezobrazena). Purifikace proteinu K2ORF7p-His provedena nebyla.

**Závěr:** V pilotním experimentu se nepodařilo ani v malém množství produkovat a purifikovat RNA polymerázu pGKL plasmidů v bakteriálním expresním systému s využitím vytvořených duálních expresních vektorů. Spolu s předchozími výsledky získanými v naší laboratoři se produkce a purifikace komplexu RNA polymerázy pGKL plasmidů, v množství umožňujícím následné biochemické experimenty *in vitro*, z klasických bakteriálních expresních systémů jeví jako nereálná.

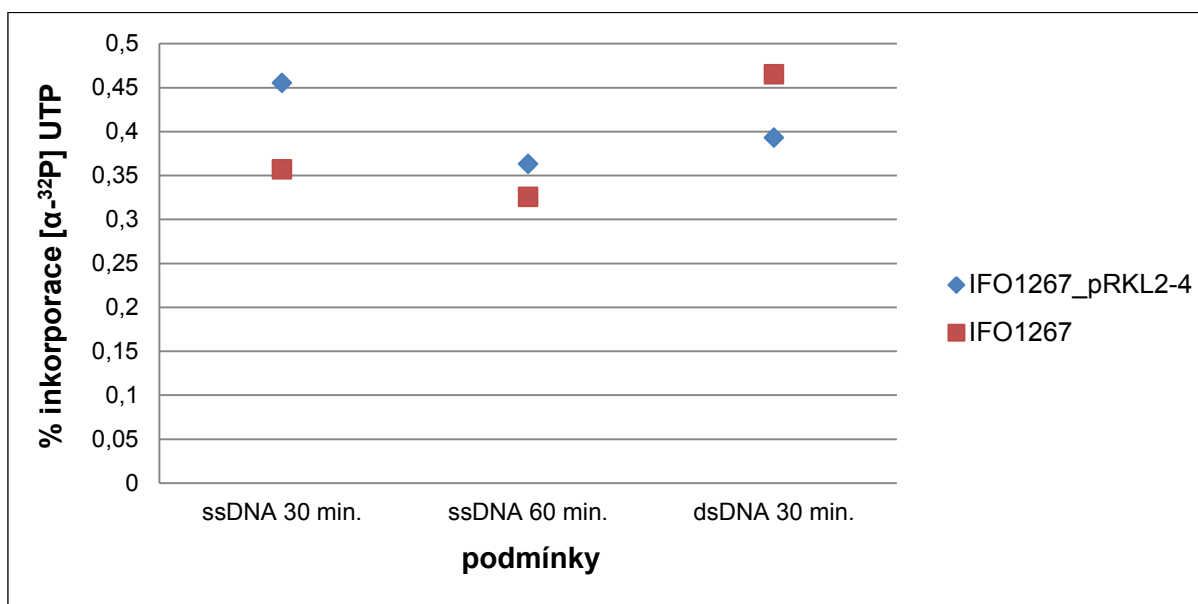
#### 5.1.4 Imunoprecipitovaný transkripční komplex pGKL plasmidů vykazuje malou či žádnou RNA polymerázovou aktivitu *in vitro*

Protože veškeré pokusy o produkci a purifikaci RNA polymerázy pGKL plasmidů v bakteriálních a kvasinkových expresních systémech v naší laboratoři selhaly (VOPÁLENSKÝ 1999; SÝKORA 2013), rozhodli jsme se pro testování RNA polymerázové aktivity předpokládané nekanonické RNAP pGKL plasmidů *in vitro* využít imunoprecipitovaný transkripční komplex z kvasinkových buněk kmene *K. lactis* IFO1267\_pRKL2-4 (EGFP-K2ORF6p). Z buněk tohoto kmene je zpravidla možné imunoprecipitovat GFP-fúzní velkou podjednotku RNAP v množství, které je v polyakrylamidových gelech detekovatelné barvením pomocí *Coomassie Blue* (Obrázek 22 a Obrázek 23) a dusičnanu stříbrného (Obrázek 24). Díky výsledkům hmotnostní spektrometrie (kapitola 5.1.2.2) a imunodetekce (kapitola 5.1.2.3) je zřejmé, že s velkou podjednotkou RNAP koprecipituje určité množství malé podjednotky RNAP (K2ORF7p), která však zpravidla není, možná i díky své malé molekulové hmotnosti, detekovatelná v polyakrylamidových gelech barvením pomocí *Coomassie Blue* nebo dusičnanu stříbrného. Pomocí metody 4.2.4.6 tak lze získat určité množství kompletní RNA polymerázy pGKL plasmidů, jejíž RNA polymerázová aktivita byla testována *in vitro* inkorporací [ $\alpha$  –  $^{32}$ P] UTP do materiálu nerozpustného v kyselině trichloroctové.

Z proteinových lyzátů kmene *K. lactis* IFO1267\_pRKL2-4 a kontrolního kmene IFO1267 byl EGFP-K2ORF6p imunoprecipitován pomocí GFP-Trap<sup>®</sup>\_A (metoda 4.2.4.6). Navázané proteiny na agarosových kuličkách byly použity pro transkripční reakci *in vitro* (metoda 4.2.3.5). Výsledný materiál byl precipitován a oddělen od [ $\alpha$  –  $^{32}$ P] UTP filtrací (metoda 4.2.3.1.5). Radioaktivita vzorků byla změřena jako počet impulzů za minutu (cpm) na kapalném scintilačním počítači (Perkin Elmer) metodou podle Čerenkova. Pro každý vzorek byla díky podílu hodnoty cpm nerozpustného materiálu k hodnotě celkového množství cpm v dané reakci vypočítána inkorporace [ $\alpha$  –  $^{32}$ P] UTP v procentech a vynesena do grafu. Naměřené hodnoty cpm jednotlivých vzorků zde nejsou uvedeny. Jako templát byla použita DNA pGKL plasmidů izolovaných z cytoplasmatické frakce lyzátu protoplastů (SÝKORA 2013), a před samotným experimentem teplotně denaturovaná, jednovláknová DNA bakteriofága M13 a nendenaturovaná dvouvláknová DNA pGKL plasmidů izolovaných z cytoplasmatické frakce lyzátu protoplastů (SÝKORA 2013). V prvním experimentu byly použity rozdílné objemy GFP-Trap<sup>®</sup>\_A s navázanými proteiny a tedy i rozdílná množství imunoprecipitovaného transkripčního komplexu pGKL plasmidů. V druhém experimentu byl použit fixní objem GFP-Trap<sup>®</sup>\_A s navázanými proteiny a rozdílné templáty a časy inkubace. Výsledky obou dvou nezávislých experimentů jsou zobrazeny na Obrázku 30 a 31.



**Obrázek 30:** Testování RNA polymerázové aktivity RNA polymerázy pGKL plasmidů. Transkripční reakce probíhala podle protokolu 4.2.3.5 po dobu 30 minut a jako templát bylo použito 1,5 μg teplotně denaturované DNA izolovaných pGKL plasmidů. Koncentrace neznačeného UTP byla 20 μM a množství značeného UTP bylo 0,21 μl [ $\alpha$  –  $^{32}$ P] UTP (3000 Ci/mmol, 10 μCi/μl).



**Obrázek 31:** Testování RNA polymerázové aktivity RNA polymerázy pGKL plasmidů. Transkripční reakce probíhala podle protokolu 4.2.3.5 s 13 μl GFP-Trap®\_A s navázanými proteiny po dobu 30 či 60 minut a jako templát bylo použito 5 μg jednovláknové DNA fága M13 (ssDNA) či 5 μg dvouvláknové DNA izolovaných pGKL plasmidů (dsDNA). Koncentrace neznačeného UTP byla 10 μM a množství značeného UTP bylo 0,28 μl [ $\alpha$  –  $^{32}$ P] UTP (3000 Ci/mmol, 10 μCi/μl).

**Závěr:** V prvním experimentu vykazovaly vzorky z kmene *K. lactis* IFO1267\_pRKL2-4 od určitého objemu GFP-Trap®\_A s navázanými proteiny mírně zvýšenou inkorporaci [ $\alpha$  –  $^{32}$ P] UTP oproti negativní kontrole, což by mohlo odpovídat slabé RNA polymerázové aktivitě imunoprecipitovaného transkripčního komplexu pGKL plasmidů. Obdobné výsledky se však nepodařilo zreprodukovat ve druhém a třetím (nezobrazeném) experimentu. Imunoprecipitovaný transkripční komplex pGKL plasmidů tak za zvolených podmínek vykazuje malou či žádnou RNA polymerázovou aktivitu *in vitro*.

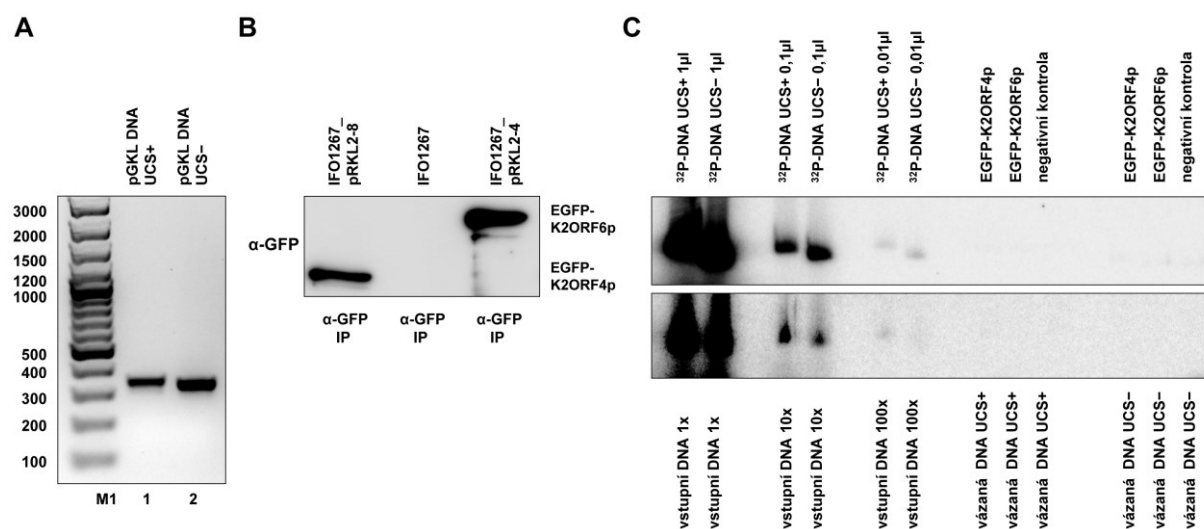
### 5.1.5 Imunoprecipitovaná RNA polymeráza a předpokládaná helikáza pGKL plasmidů nevykazuje DNA vazebnou aktivitu *in vitro*

Pokusy o produkci a purifikaci předpokládané helikázy pGKL plasmidů v bakteriálních a kvasinkových expresních systémech v naší laboratoři také nebyly příliš úspěšné. Produkce His-K2ORF4p a GST-K2ORF4p z vektorů pIVEX2.4c/ORF4 a pGEX4T2/ORF4 v bakteriálních expresních systémech využívající různé BL21(DE3)-odvozené kmeny *E. coli* nebyla detekována (PEJZNOCHOVÁ 2004). Slabá produkce HA-K2ORF4p z vektoru pYX213/ORF4 v kmeni *S. cerevisiae* HW303-1A sice byla zaznamenána (PEJZNOCHOVÁ 2004), ale opakováním stejného postupu se mi tento protein produkovat nepodařilo (data nezobrazena). Stejně jako v případě velké podjednotky RNAP je však možné imunoprecipitovat GFP-fúzní předpokládanou helikázu z příslušného kmene (*K. lactis* IFO1267\_pRKL2-8) v množství, které je v polyakrylamidových gelech detekovatelné barvením pomocí *Coomassie Blue* (data nezobrazena). Imunoprecipitovaná helikáza, stejně jako imunoprecipitovaný transkripční komplex, byly použity k DNA vazebným pokusům *in vitro*.

Z proteinových lyzátů kmene *K. lactis* IFO1267\_pRKL2-4 (EGFP-K2ORF6p), IFO1267\_pRKL2-8 (EGFP-K2ORF4p) a kontrolního kmene IFO1267 byly GFP-fúzní proteiny imunoprecipitovány pomocí GFP-Trap®\_A (metoda 4.2.4.6). Navázané proteiny na agarosových kuličkách byly použity pro experiment testující vazbu radioaktivně značené DNA *in vitro* (4.2.2.1.11), adaptovaný podle (KIM 2004). Krátký úsek DNA pGKL plasmidů neobsahující žádnou promotorovou sekvenci UCS byl amplifikován pomocí PCR (metody 4.2.2.6.1 a 4.2.2.6.3) s primery in\_ORF5\_forw a in\_ORF5\_rev, rozdělen agarosovou elektroforézou (metoda 4.2.2.3.1) a izolován z gelu (metoda 4.2.2.1.5). Obdobně byl připraven krátký úsek DNA pGKL plasmidů obsahující dvě promotorové sekvence UCS PCR amplifikací s primery 5RACE\_O8\_K2 a in\_ORF9\_rev. Elektroforetogramy připravených a izolovaných DNA fragmentů jsou na Obrázku 32A. Oba DNA fragmenty byly radioaktivně značeny fosforylací 5' OH skupin (metoda 4.2.2.4.7), přečištěny od [ $\gamma$  –  $^{32}$ P] ATP (metoda 4.2.2.1.9) a použity



k testování vazby DNA imunoprecipitovanými proteiny *in vitro*. Navázaná DNA byla analyzována pomocí polyakrylamidové elektroforézy (metoda 4.2.2.3.3), přítomnost imunoprecipitovaných proteinů ve vzorcích byla potvrzena pomocí SDS-PAGE (metoda 4.2.4.8) a imunodetekce pomocí metody *western blot* (metody 4.2.4.9 a 4.2.4.10) viz Obrázek 32B. Značení DNA, vazba DNA *in vitro* a analýza vázané DNA proběhla v Laboratoři mikrobiální genetiky a genové exprese Libora Krásného za asistence Petry Sudzinové. Pokus byl proveden ve dvou biologických replikátech a výsledky jsou zobrazeny na Obrázku 32C.

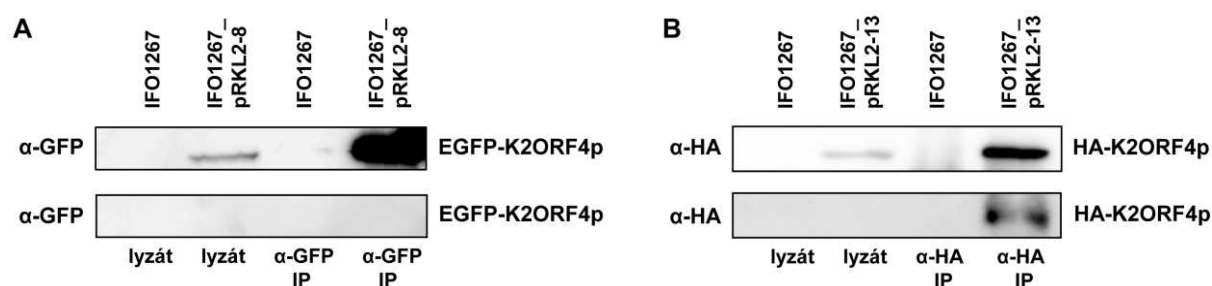


**Obrázek 32:** Testování DNA vazebné aktivity předpokládané helicázy a RNA polymerázy pGKL plasmidů. (A) Elektroforetogram dokumentující přípravu fragmentů DNA pGKL plasmidů obsahujících (UCS+) a neobsahujících (UCS-) promotorový element UCS pGKL plasmidů. M1: GeneRuler™ 100 bp Plus DNA Ladder; 1: PCR fragment plasmidu pGKL2 připravený za použití primeru 5RACE\_O8\_K2 a in\_ORF9\_rev (produkt 350 bp); 2: PCR fragment plasmidu pGKL2 připravený za použití primeru in\_ORF5\_forw a in\_ORF5\_rev (produkt 331 bp). (B) *Western blot* analýza imunoprecipitovaných proteinů. EGFP-K2ORF4p byl purifikován pomocí GFP Trap®\_A (α-GFP IP) z lyzátů buněk *K. lactis* IFO1267\_pRKL2-8; EGFP-K2ORF6p byl purifikován z lyzátů buněk IFO1267\_pRKL2-4. Purifikované proteiny z obou kmenů a kontrolního kmene IFO2167 byly použity pro testování vazby radioaktivně značené DNA *in vitro*. Pro imunodetekci proteinů byla použita primární protilátka α-GFP B-2 (1:1000) a sekundární protilátka α-mouse IgG HRP (1:5000). (C) Radioelektroforetogram dvou nezávislých DNA vazebných experimentů *in vitro*. Zobrazena je 1/50 objemu 1x, 10x a 100x ředěné vstupní DNA a 1/2 objemu eluované DNA, vázané imunoprecipitovanými proteiny z kmenů IFO1267\_pRKL2-8 (EGFP-K2ORF4p), IFO1267\_pRKL2-4 (EGFP-K2ORF6p) a IFO1267 (negativní kontrola).

**Závěr:** Ve dvou nezávislých experimentech se nepodařilo detekovat vazbu <sup>32</sup>P-značené DNA obsahující i neobsahující promotorový element pGKL plasmidů na imunoprecipitovanou předpokládanou helicázu ani na imunoprecipitovaný transkripční komplex pGKL plasmidů. Tyto imunoprecipitované proteiny tak za zvolených podmínek nevykazují DNA vazebnou aktivitu *in vitro*.

### 5.1.6 Předpokládaná helikáza pGKL plasmidů vykazuje specifickou DNA vazebnou aktivitu *in vivo*

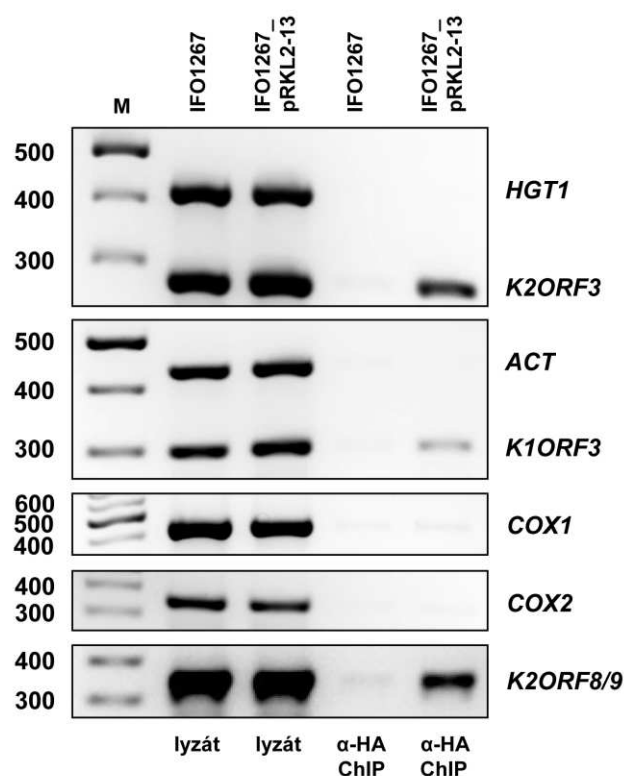
Předpokládaná helikáza pGKL plasmidů K2ORF4p je sekvenčně podobná a fylogeneticky příbuzná s transkripčními helikázami D6 a NPH I viru *Vaccinia* (WILSON a MEACOCK 1988; IYER *et al.* 2001), a oba tyto proteiny vykazují DNA vazebnou aktivitu *in vitro* (PAOLETTI a MOSS 1974; LI a BROYLES 1993a). Proto jsem se po neúspěšných pokusech *in vitro* rozhodl testovat vazbu K2ORF4p na DNA pGKL plasmidů *in vivo* pomocí chromatinové imunoprecipitace (ChIP). K pokusům byly použity připravené kmeny *K. lactis* IFO1267\_pRK2-8 (EGFP-K2ORF4p), IFO1267\_pRK2-13 (HA-K2ORF4p). *Crosslink* byl proveden pomocí formaldehydu. Protein EGFP-K2ORF4p lze pomocí imunoprecipitace využívající GFP-Trap®\_A izolovat v množství, které je v polyakrylamidových gelech detekovatelné barvením pomocí *Coomassie Blue*, a také imunodetekovatelné metodou *western blot* (Obrázek 33A). Protein HA-K2ORF4p lze pomocí imunoprecipitace využívající anti-HA HA-7 agarosu izolovat v množství, které je imunodetekovatelné metodou *western blot* (Obrázek 33B). Po opůsobení buněk příslušných kmenů formaldehydem a následné imunoprecipitaci rekombinantních proteinů z lyzátů však protein EGFP-K2ORF4p v imunoprecipitátu již není imunodetekovatelný (Obrázek 33A), ale protein HA-K2ORF4p zůstává v imunoprecipitátu imunodetekovatelný, a tím pádem jej prokazatelně lze využít pro účely chromatinové imunoprecipitace (Obrázek 33B). Pro účely chromatinové imunoprecipitace byl proto dále používán pouze kmen *K. lactis* IFO1267\_pRK2-13.



**Obrázek 33:** *Western blot* analýza imunoprecipitované předpokládané helikázy pGKL plasmidů. (A) EGFP-K2ORF4p byl purifikován pomocí GFP Trap®\_A (α-GFP IP) z lyzátů buněk *K. lactis* IFO1267\_pRK2-8 a kontrolního kmene IFO1267 (horní panel), a z lyzátů buněk stejných kmenů, které byly před lyzí opůsobeny formaldehydem (spodní panel). (B) HA-K2ORF4p byl purifikován pomocí anti-HA HA-7 agarosu (α-HA IP) z lyzátů buněk *K. lactis* IFO1267\_pRK2-13 a kontrolního kmene IFO1267 (horní panel), a z lyzátů buněk stejných kmenů, které byly před lyzí opůsobeny formaldehydem (spodní panel). Pro imunodetekci proteinů byla použita primární protilátka α-GFP B-2 (ředění 1:1000), α-HA 6E2 (1:1000) a sekundární protilátka α-mouse IgG HRP (1:5000). Lyzát představoval přibližně 1/100 vzorku a IP představoval přibližně 1/2 vzorku.

Z proteinových lyzátů buněk kmenů *K. lactis* IFO1267\_pRKL2-13 (HA-K2ORF4p) a kontrolního kmene IFO1267 opůsobených formaldehydem byla imunoprecipitována chromatinová frakce podle protokolu 4.2.2.1.12. O sestavení vhodného protokolu a první fungující chromatinovou imunoprecipitaci na použitých kmenech se zasloužil kolega Josef Novák. Vstupní a imunoprecipitovaná DNA byla analyzována pomocí PCR a následně agarosové elektroforézy. Geny a primery použité k analýze pomocí PCR jsou uvedeny v Tabulce 17. Ve dvou biologických replikátech (zobrazeny výsledky pouze z prvního experimentu) byla potvrzena specifická interakce předpokládané helikázy K2ORF4p s DNA pGKL plasmidů *in vivo*, a tato helikáza *in vivo* neinteraguje s chromosomální ani mitochondriální DNA hostitelských buněk (Obrázek 34).

Tabulka 17: Analýza chromatinové imunoprecipitace K2ORF4p pomocí PCR				
Gen	Umístění	Primery	Produkt	Počet cyklů
ACT	chromosom D	actin_KL-RNAfor + aktin_KL-rev	436 bp	25
COX1	mitochondriální DNA	Sc/Kl_Cox1-for + Sc/Kl_Cox1-rev	426 bp	25
COX2	mitochondriální DNA	Sc/Kl_Cox2-for + Sc/Kl_Cox2-rev	349 bp	25
HGT1	chromosom A	HGT1_KL-forw + HGT1_KL-rev	420 bp	25
K1ORF3	plasmid pGKL1	K05 + 5RACE_O3_K1_3	304 bp	25
K2ORF3	plasmid pGKL2	vORF3-k2-forw + vORF3-k2-rev_2	276 bp	25
K2ORF8/9	plasmid pGKL2	5_RACE_O8_K2 + in_ORF9_rev	350 bp	25



**Obrázek 34:** Fyzické spojení předpokládané helikázy K2ORF4p pGKL plasmidů s plasmidovou DNA. Elektroforetogram PCR analýzy přítomnosti chromosomální DNA (*ACT*, *HGT1*), mitochondriální DNA (*COX1*, *COX2*) a DNA pGKL plasmidů (*K1ORF3*, *K2ORF3*, *K2ORF8/9*) v sonikovaných lyzátech a HA-imunoprecipitovaném chromatinu z buněk kmene *K. lactis* IFO1267\_pRKL2-13 (HA-K2ORF4p) a kontrolního kmene IFO1267. Některé vzorky individuálně provedených specifických PCR reakcí byly před elektroforetickou analýzou sloučeny. Identita fragmentů jednotlivých genů je uvedena vpravo. PCR amplifikace byly provedeny na 1/1200 objemu chromatinu před imunoprecipitací (lyzát) a 1/30 objemu chromatinové imunoprecipitace (α-HA ChIP) za použití primerů uvedených v Tabulce 17. M: GeneRuler™ 100 bp Plus DNA Ladder.

### 5.1.6.1 Kvantifikace chromatinové imunoprecipitace pomocí real-time PCR

Ačkoliv je asociace plasmidové DNA s proteinem HA-K2ORF4p zřejmá pouze z jednoduché PCR analýzy a následného elektroforetického dělení, rozhodl jsem se asociaci plasmidové DNA s předpokládanou helikázou *in vivo* kvantifikovat pomocí citlivější semikvantitativní real-time PCR metody. S návrhem, provedením a vyhodnocením real-time PCR experimentů asistovala kolegyně Silvie Mrvová.

Vzorky odpovídající 1/2000 objemu sonikovaných lyzátů a 1/30 objemu chromatinových imunoprecipitátů byly 10x naředěny a 2,5 µl bylo použito jako templát pro semikvantitativní real-time PCR amplifikaci v triplikátu využívající SYBR Green I (metody 4.2.2.7.1 a 4.2.2.7.2) a specifické primery viz Tabulka 17. Vzorek sonikovaného lyzátu kmene IFO1267 byl v pěti desítkových ředěních a triplikátu použit pro stanovení standardní křivky každého páru použitých primerů a vypočtení účinnosti dané amplifikace. Relativní kvantifikace real-time PCR a kontrola křivek tání výsledných produktů probíhala v programu LightCycler® 480 Software, Verze 1.5 (Roche). Část vzorků byla analyzována na přítomnost specifických PCR produktů po proběhnutí real-time PCR pomocí agarosové elektroforézy. Bohužel většina párů primerů se v použitém experimentálním uspořádání ukázala být pro účely real-time PCR nevhodná, protože reakce buď probíhala s malou účinností (např. *K2ORF8/9*) nebo specifický produkt vznikl minoritně či nevznikal vůbec (např. *COX1*, *COX2*, *K1ORF3*; data nezobrazena). Aby byl v analýze zastoupen alespoň jeden gen kódovaný mitochondriální DNA, byly navrženy, a pro amplifikaci fragmentu *COX1* genu úspěšně použity, další primery Kl\_Cox1-for2 a Kl\_Cox1-rev2. Hodnoty standardních křivek amplikonů vybraných genů ve dvou nezávislých biologických experimentech jsou uvedeny v Tabulce 18. Průměrné hodnoty prahového cyklu (Ct) vybraných amplikonů každého vzorku obou pokusů jsou uvedeny v Tabulce 19.

**Tabulka 18:** Hodnoty standardních křivek a účinnosti jednotlivých amplifikací

Gen	Směrnice	Průsečík s y	Účinnost
Pokus 1			
<i>ACT</i>	-3,468	13,03	1,943
<i>K2ORF3</i>	-3,579	9,334	1,903
<i>COX1</i>	-4,257	6,526	1,718
Pokus 2			
<i>ACT</i>	-3,639	16,10	1,883
<i>K2ORF3</i>	-2,725	12,95	2,328
<i>COX1</i>	-3,387	10,05	1,973

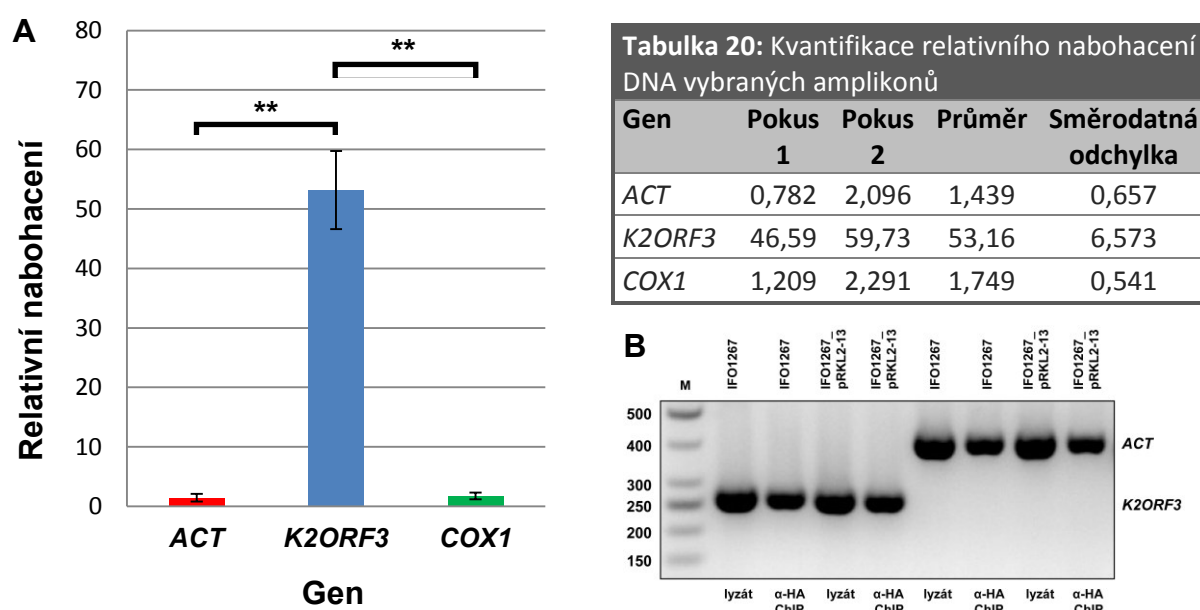
**Tabulka 19:** Průměrné hodnoty Ct jednotlivých amplikonů

Gen	IFO1267 lyzát	IFO1267_ pRKL2-13 lyzát	IFO1267 ChIP	IFO1267_ pRKL2-13 ChIP
Pokus 1				
<i>ACT</i>	19,28	20,31	26,67	28,07
<i>K2ORF3</i>	14,11	15,92	24,98	20,82
<i>COX1</i>	14,61	14,23	24,20	23,47
Pokus 2				
<i>ACT</i>	21,47	23,04	31,32	31,72
<i>K2ORF3</i>	20,30	21,31	29,47	25,64
<i>COX1</i>	14,57	16,56	25,18	25,95

Relativní kvantifikace nabohacení DNA ampliconů příslušných genů ve vzorku ChIP kmene *K. lactis* IFO1267\_pRK2-13 probíhala s využitím průměrných hodnot Ct a účinnosti ( $e$ ) jednotlivých amplifikačních reakcí daného experimentu podle následujícího vzorce, modifikovaného podle (PASCUAL-AHUIR a PROFT 2012):

$$\text{Relativní nabohacení} = e^{(Ct_{\text{lyzát}} - Ct_{\text{ChIP}})_{\text{IFO1267\_pRK2-13}} - (Ct_{\text{lyzát}} - Ct_{\text{ChIP}})_{\text{IFO1267}}}$$

Vypočítané hodnoty relativního nabohacení DNA jednotlivých genů z obou experimentů, jejich aritmetický průměr a směrodatná odchylka jsou uvedeny v Tabulce 20. Dvě nezávislá opakování experimentu neposkytovala prostor pro statistické vyhodnocení o vysoké výpovědní hodnotě, přesto jsem se alespoň pokusil o vzájemné porovnání průměrů a směrodatných odchylek relativních nabohacení ampliconů s využitím volně dostupného internetového nástroje<sup>3</sup>. V použitém porovnání vyšel rozdíl průměrů relativního nabohacení ampliconu pGKL plasmidu oproti ampliconu chromosomální či mitochondriální DNA ve vzorku ChIP IFO1267\_pRK2-13 jako signifikantní s hodnotou  $P < 0,01$  při intervalu spolehlivosti 95 % (Obrázek 35A). Ověření identity ampliconů u některých vybraných vzorků po proběhnutí real-time PCR pomocí agarosové elektroforézy je zobrazeno na Obrázku 35B.



**Obrázek 35:** Fyzické spojení předpokládané helikázy K2ORF4p pGKL plasmidů s plasmidovou DNA, kvantifikované pomocí semikvantitativní real-time PCR. **(A)** Sloupcový graf průměrných hodnot a směrodatných odchylek relativního nabohacení DNA ampliconu chromosomální DNA (*ACT*, červeně), plasmidové DNA (*K2ORF3*, modře) a mitochondriální DNA (*COX1*, zeleně) v chromatinu imunoprecipitovaném z buněk kmene *K. lactis* IFO1267\_pRK2-13 ze dvou nezávislých experimentů. \*\*: hladina signifikance  $P < 0,01$ . **(B)** Elektroforetogram ověření identity ampliconů vybraných vzorků po proběhnutí real-time PCR (45 cyklů). Identita fragmentů je uvedena vpravo. Real-time PCR amplifikace byly provedeny na 1/4800 objemu chromatinu před imunoprecipitací (lyzát) a 1/120 objemu chromatinové imunoprecipitace ( $\alpha$ -HA ChIP) za použití primerů uvedených v Tabulce 17. M: GeneRuler™ 50 bp Plus DNA Ladder.

<sup>3</sup> [http://www.medcalc.org/calc/comparison\\_of\\_means.php](http://www.medcalc.org/calc/comparison_of_means.php)



**Závěr:** Výsledky dvou nezávislých experimentů chromatinové imunoprecipitace proteinu HA-K2ORF4p analyzované pomocí PCR v kombinaci s agarosovou elektroforézou a pomocí semikvantitativní real-time PCR jasně ukazují, že předpokládaná helikáza pGKL plasmidů asociuje *in vivo* s plasmidovou DNA. Interakce předpokládané helikázy s plasmidovou DNA je specifická, protože předpokládaná helikáza neinteraguje s chromosomální ani mitochondriální DNA hostitelských buněk *in vivo*. Vzhledem k predikovaným vlastnostem proteinu K2ORF4p bude zřejmě interakce předpokládané helikázy s plasmidovou DNA přímá, ačkoliv se nedá vyloučit interakce zprostředkovaná, například přes některý protein transkripčního komplexu pGKL plasmidů.

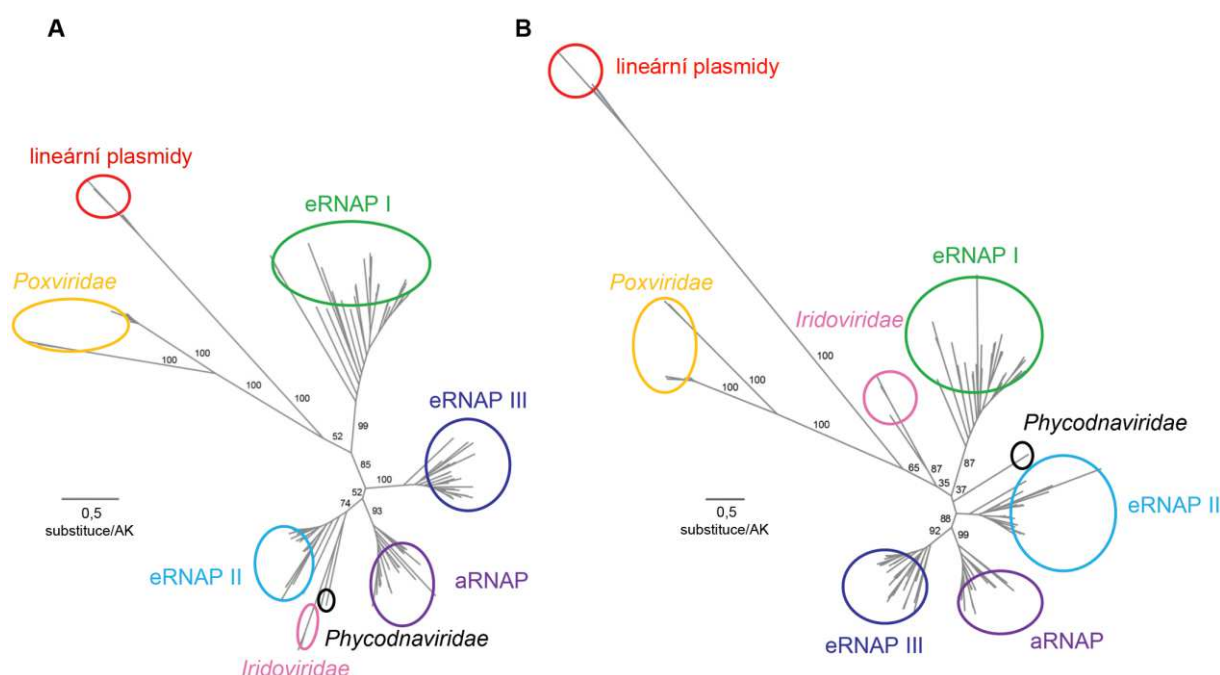
### 5.1.7 RNA polymeráza pGKL plasmidů má poxvirový původ

Rozsáhlá fylogenetická analýza proteinových sekvencí RNA polymeráz kvasinkových lineárních plasmidů dosud nebyla publikována. Dřívější analýzy byly zaměřeny pouze na konzervované oblasti  $\beta'$ a11 a  $\beta'$ a12 proteinu K2ORF6p. Tyto studie naznačovaly, že plasmidová RNAP má blíže k vícepodjednotkovým RNA polymerázám, než k jednopodjednotkovým RNAP, kódovaným mitochondriálními lineárními plasmidy rostlin a hub (KEMPKE *et al.* 1992; ROHE *et al.* 1992). Proto se předpokládá, že geny *ORF6* a *ORF7* kvasinkových lineárních plasmidů byly odvozeny z eukaryotických genů kódujících RNAP hostitelských buněk (JESKE *et al.* 2007; RUPRICH-ROBERT a THURIAUX 2010) nebo alternativně, že gen *ORF6* je dávným zástupcem diverzifikace vícepodjednotkových RNAP z časů, kdy se  $\beta$  a  $\beta'$  podjednotky nacházely na jednom proteinu (IYER a ARAVIND 2012).

Již ve své diplomové práci jsem provedl podrobnou fylogenetickou analýzu, která by osvětlila evoluční původ RNAP kvasinkových lineárních plasmidů (SÝKORA 2013). Použil jsem sekvence konzervovaných oblastí  $\beta$ a1,  $\beta$ a3– $\beta$ a4,  $\beta$ a6– $\beta$ a8,  $\beta$ a10– $\beta$ a16,  $\beta'$ a11– $\beta'$ a12,  $\beta'$ a15– $\beta'$ a16 a  $\beta'$ a20 proteinů *ORF6p* a *ORF7p* ze všech dosud sekvenovaných kvasinkových cytoplasmatických lineárních plasmidů (rozložení konzervovaných oblastí RNAP v rámci proteinu K2ORF6p a K2ORF7p je znázorněno na Obrázku 2). Tyto konzervované oblasti byly rozděleny do dvou sekvenčních porovnání, které odpovídaly oblastem  $\beta$  a  $\beta'$  podjednotky. Tato sekvenční porovnání byla potom spojena s publikovanými sekvenčními porovnáními konzervovaných oblastí  $\beta$  a  $\beta'$  podjednotky kanonických vícepodjednotkových RNAP (LANE a DARST 2010a) a následně použita pro konstrukci fylogenetických stromů, viz (SÝKORA 2013). Pro účely disertační práce jsem se rozhodl na základě svých sekvenčních porovnání zkonstruovat obdobné fylogenetické stromy, které by obsahovaly informaci o podpoře jednotlivých větví stromu na základě *bootstrap* analýzy. Fylogenetické stromy byly vytvořeny v programu PhyML 3.0



a vizualizovány v programu Rainbow Tree (kapitola 4.2.7). Obrázek 36A a B znázorňuje nezakořeněné fylogenetické stromy porovnání konzervovaných oblastí  $\beta$  a  $\beta'$  podjednotek archeálních, eukaryotických, virových a plasmidových RNA polymeráz (230 sekvencí). Stromy ukazují, že plasmidové RNAP se větví společně s virovými RNAP z rodiny *Poxviridae*, a konzervované oblasti  $\beta'$  podjednotky plasmidových RNAP se také větví blízko virovým RNAP z rodiny *Iridoviridae* (Obrázek 36B). To je v souladu s převládajícím názorem, že původní ortology  $\beta'$  podjednotky všech nukleocytoplasmových virů jsou monofyletické a odvozené z eukaryotické RNA polymerázy I (YUTIN a KOONIN 2012).



**Obrázek 36:** Fylogenetická analýza původu RNA polymeráz kvasinkových lineárních plasmidů. *Maximum likelihood* fylogenetické stromy byly zkonstruovány v PhyML 3.0. Fylogenetické stromy jsou zobrazeny jako nezakořeněné fylogramy, kde je délka jednotlivých větví stromu zobrazena proporciálně s vypočítanou evoluční vzdáleností jednotlivých sekvencí. Okraje listů ohraničující jednotlivé typy RNA polymeráz jsou barevně odlišeny. Použité zkratky pro skupiny RNAP: aRNAP, archeální RNAP; eRNAP, eukaryotické RNAP. Pro oba zobrazené stromy je uvedeno měřítko délky větví v průměrné hodnotě 0,5 substitucí na aminokyselinový zbytek. Vybrané hodnoty podpory větví vypočítané z 1000 *bootstrap* opakování jsou označeny černě. **(A)** Fylogram ortologů  $\beta$  podjednotek na základě sekvenčního porovnání aminokyselinových sekvencí všech konzervovaných oblastí  $\beta$  podjednotek vybraných kanonických RNAP a těch konzervovaných oblastí  $\beta$  podjednotky, přítomných v genech ORF6 kvasinkových cytoplasmatických lineárních plasmidů. **(B)** Fylogram ortologů  $\beta'$  podjednotek na základě sekvenčního porovnání aminokyselinových sekvencí všech konzervovaných oblastí  $\beta'$  podjednotek vybraných kanonických RNAP a těch konzervovaných oblastí  $\beta'$  podjednotky, přítomných v genech ORF6 a ORF7 kvasinkových cytoplasmatických lineárních plasmidů. Podrobnosti viz kapitola 4.2.7 a SÝKORA (2013).

**Závěr:** Fylogenetická analýza překvapivě poukazuje na virový původ RNA polymeráz kvasinkových lineárních plasmidů, blízky k virům čeledi *Poxviridae*.

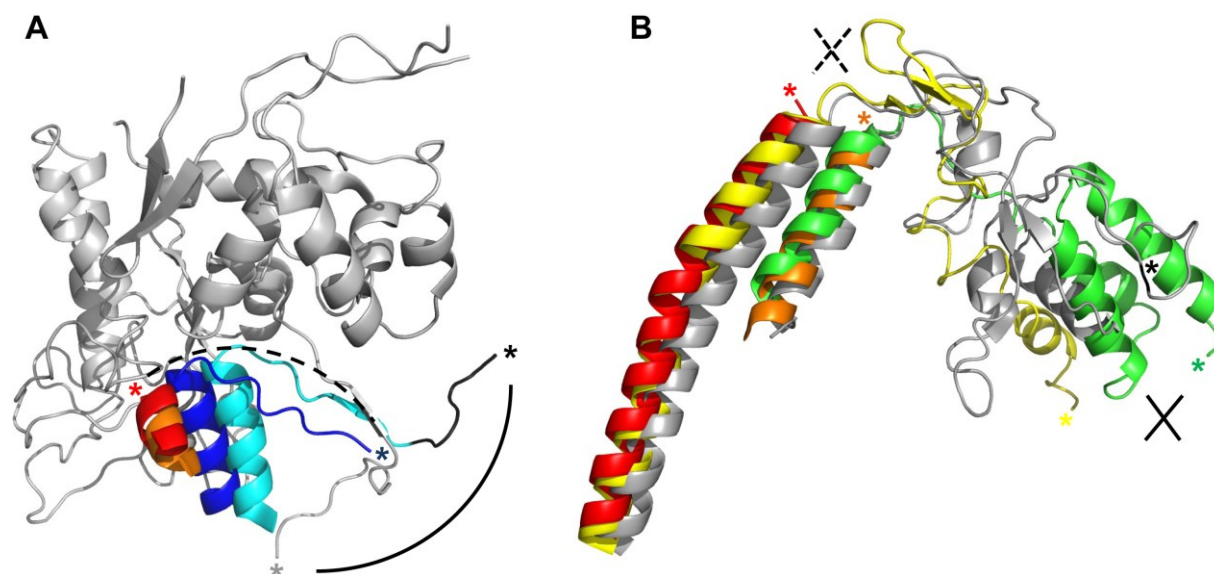
### 5.1.8 RNA polymeráza pGKL plasmidů má unikátní architekturu

Jak již bylo zmíněno v kapitole 2.1.1.1, RNA polymeráza pGKL plasmidů patří mezi nekanonické RNAP vykazující sekvenční podobnost pouze k  $\beta$  a  $\beta'$  podjednotkám vícepodjednotkových RNAP. Do této skupiny nekanonických RNA polymeráz patří ještě RNAP eukaryotických virů skupiny *Baculoviridae*, některých bakteriofágů a některých bakteriálních mobilních elementů [pro přehled viz (SÝKORA 2013)]. V rámci méj diplomové práce jsem vytvořil model terciární struktury větší části RNAP kódované pGKL plasmidy (SÝKORA 2013). Z vytvořeného modelu a dřívějších sekvenčních porovnání je patrná značná redukce RNAP lineárních plasmidů oproti kanonickým RNAP, zejména na solventní straně enzymu a oblastech umístěných distálně od aktivního centra enzymu. RNAP lineárních plasmidů má však kromě značné redukce také unikátní architekturu nevyskytující se u žádných jiných, ať už kanonických či nekanonických, vícepodjednotkových RNA polymeráz.

K této unikátní architektuře přispívají dva strukturní prvky. Prvním z nich je spojení  $\beta$  podjednotky s částí  $\beta'$  podjednotky u proteinu K2ORF6p (Obrázek 2). Jak již bylo zmíněno v kapitole 2.1.1.1, fúze  $\beta$  a  $\beta'$  podjednotky RNAP je běžná u některých parazitických bakterií, například u rodu *Helicobacter* a *Wolbachia*. Zde se ale vždy se jedná o fúzi mezi konzervovanou oblastí  $\beta$ a16 a  $\beta'$ a1 (LANE a DARST 2010a). U *Helicobacter pylori* je možné uměle oddělit  $\beta$  a  $\beta'$  podjednotku RNAP, přičemž je výsledný kmen životaschopný v buněčné kultuře i *in vivo*, a vykazuje pouze citlivost na působení močoviny (DAILIDIENE *et al.* 2007). Je také možná umělá fúze mezi oblastí  $\beta$ a16 a  $\beta'$ a1 u  $\beta$  a  $\beta'$  podjednotek RNA polymerázy *E. coli*, přičemž fúzní  $\beta$ : $\beta'$  protein se skládá ve funkční RNAP *in vitro* a *in vivo* (SEVERINOV *et al.* 1997). U proteinu K2ORF6p je spojena konzervovaná oblast  $\beta$ a16 s oblastí  $\beta'$ a11 (SÝKORA 2013), což u kanonických RNAP odpovídá spojení  $\beta$  podjednotky pouze s druhou polovinou  $\beta'$  podjednotky, přičemž podobná fúze nebyla u žádné jiné RNAP popsána (Obrázek 37A).

Druhým charakteristickým strukturním prvkem RNAP lineárních plasmidů je rozdělení  $\beta'$  podjednotky na dva proteiny (K2ORF6p a K2ORF7p) mezi konzervovanou oblastí  $\beta'$ a15 a  $\beta'$ a16 (SCHAFFRATH *et al.* 1995a; SÝKORA 2013). Obdobné rozdělení  $\beta'$  podjednotky na dva polypeptidy mezi konzervovanou oblastí  $\beta'$ a15 a  $\beta'$ a16 je známo u kanonických RNA polymeráz většiny druhů *Archaea* (KWAPISZ *et al.* 2008). Je možné uměle rozdělit  $\beta'$  podjednotku RNA polymerázy *E. coli* mezi oblastmi  $\beta'$ a15 a  $\beta'$ a16 obdobně jako u *Archaea*, přičemž rozdělené proteiny se skládají ve funkční RNAP *in vitro* (SEVERINOV *et al.* 1996). Také je možné rozdělené  $\beta'$  podjednotky archeální RNA polymerázy uměle spojit při současném zachování funkcí proteinu (WERNER a GROHMANN 2011). Mezi oblastmi  $\beta'$ a15 a  $\beta'$ a16 se u kanonických RNAP nachází insert, který není mezi bakteriálními a archeo-eukaryotickými RNAP konzervovaný

(IYER *et al.* 2003; LANE a DARST 2010a). K přirozenému rozdělení  $\beta'$  podjednotky archeálních RNA polymeráz a umělému rozdělení  $\beta'$  podjednotky RNAP *E. coli* dochází v rámci tohoto insertu, který však RNAP pGKL plasmidů úplně postrádá viz Obrázek 37B (SÝKORA 2013).



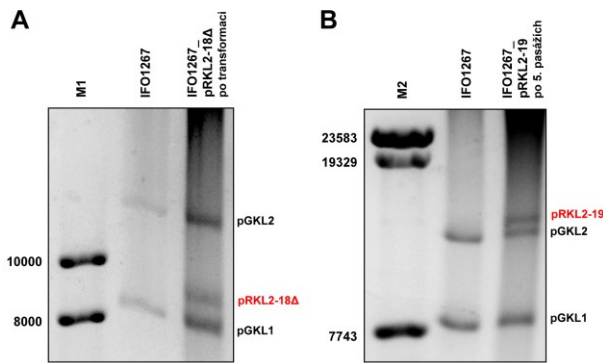
**Obrázek 37:** Unikátní architektura RNA polymerázy pGKL plasmidů. **(A)** Oblasti fúze  $\beta$  podjednotky s  $\beta'$  podjednotkou. Zobrazena je část konzervované oblasti  $\beta$ a16 modelu  $\beta$  úseku K2ORF6p (modře, C-konec označen modrou hvězdičkou) a homologní oblast struktury proteinu Rpb2 RNAP II *Saccharomyces cerevisiae* (PDB ID: 2NVQ, tyrkysově, nehomologní část označena černě, C-konec označen černou hvězdičkou). U kanonických RNAP může být někdy spojen C-konec  $\beta$  podjednotky s blízkým N-koncem  $\beta'$  podjednotky (označen šedou hvězdičkou). Umístění linkeru fúze je naznačeno čarou. U RNAP pGKL plasmidů je spojen C-konec  $\beta$  podjednotky s oblastí  $\beta'$ a11. Umístění linkeru fúze je naznačeno přerušovanou čarou. Zobrazena je část oblasti  $\beta'$ a11 modelu  $\beta'$  úseku K2ORF6p (červeně, N-konec označen červenou hvězdičkou) a homologní oblast struktury proteinu Rpb1 RNAP II *Saccharomyces cerevisiae* (PDB ID: 2NVQ, oranžově). Velká N-koncová část  $\beta'$  podjednotky, která RNA polymeráze pGKL plasmidů chybí, je zobrazena na struktuře proteinu Rpb1 RNAP II *Saccharomyces cerevisiae* (PDB ID: 2NVQ, šedě). **(B)** Oblasti rozdělení  $\beta'$  podjednotky na dva polypeptidy u *Archaea* a lineárních plasmidů. Zobrazena je část konzervované oblasti  $\beta'$ a15 modelu  $\beta'$  úseku K2ORF6p (červeně) po C-konec proteinu (červená hvězdička) a část konzervované oblasti  $\beta'$ a16 modelu K2ORF7p (oranžově) od N-konce proteinu (oranžová hvězdička). Umístění rozdělení je naznačeno přerušovaným křížkem. Dále je zobrazena část oblasti  $\beta'$ a15 s insertem proteinu Rpo1N RNAP *Sulfolobus shibatae* (PDB ID: 2WAQ, žlutě) po C-konec proteinu (žlutá hvězdička) a část oblasti  $\beta'$ a16 s insertem proteinu Rpo1C RNAP *Sulfolobus shibatae* (PDB ID: 2WAQ, zeleně) od N-konce proteinu (zelená hvězdička). Umístění rozdělení je naznačeno křížkem. Znázorněna je též část struktury  $\beta'$  podjednotky RNAP *Escherichia coli* (PDB ID: 4LK1, šedě) a pozice místa, ve kterém došlo k umělému rozdělení tohoto proteinu na dva polypeptidy v práci SEVERINOV *et al.* (1996), je zobrazena černou hvězdičkou. Modely byly se strukturami porovnány pomocí DaliLite v. 3 a vizualizovány pomocí PyMOL 1.3.

Pokusil jsem se testovat vliv změny charakteristických prvků unikátní architektury RNAP pGKL plasmidů na její funkci *in vivo* rozdělením proteinu K2ORF6p na dvě části odpovídající  $\beta$  a  $\beta'$  podjednotce, a také fúzí proteinů K2ORF6p a K2ORF7p mezi konzervovanými oblastmi  $\beta'$ a15 a  $\beta'$ a16. Byly připravovány dva kmeny *K. lactis* s modifikovanými pGKL plasmidy

postupem detailně popsaným v kapitole 5.1.1.1. Kmen IFO1267\_pRKL2-19 měl obsahovat modifikovaný plasmid pGKL2 kódující místo genu *K2ORF6* (1–974) geny *K2ORF6β* (1–693) a *K2ORF6β'* (694–974), oba pod kontrolou přirozeného promotoru K2UCR6. Kmen IFO1267\_pRKL2-18 měl obsahovat modifikovaný plasmid pGKL2 kódující místo genů *K2ORF6* a *K2ORF7* fúzní gen *K2ORF6-7* pod kontrolou K2UCR6. Ke spojení β'a15 a β'a16 konzervovaných oblastí proteinů K2ORF6p a K2ORF7p byl použit insert, nacházející se mezi konzervovanými oblastmi β'a15 a β'a16 proteinu Rpo147 viru *Vaccinia* kmene Western Reserve (aminokyselinové zbytky 765–924), protože se jedná o fylogeneticky příbuzný protein RNAP pGKL plasmidů. Nejdůležitější informace k přípravě a charakterizaci nových kmenů souvisejících s touto kapitolou jsou shrnuty v Tabulce 21. Kmen IFO1267\_pRKL2-19 se podařilo připravit, avšak rekombinantní plasmid pRKL2-19 nedokázal v buňkách nahradit plasmid pGKL2 při inkubaci v selektivních podmínkách favorizujících modifikovaný plasmid (Obrázek 38B). Kmen IFO1267\_pRKL2-18 se navzdory pěti nezávislým pokusům o transformaci připravit nepodařilo, avšak podařilo se připravit kmen, který obsahoval zkrácený rekombinantní plasmid pRKL2-18Δ, obsahující část integrační kazety (Obrázek 38A).

**Tabulka 21:** Primery a templáty pro přípravu a charakterizaci kmenů této kapitoly

Plasmid	Primery na přípravu kazety pomocí PCR	Templát pro přípravu kazety	Primery na přípravu kazety pomocí fúzní PCR	Primery pro přípravu templátu na sekvenování / Sekvence	Primery pro ověření integrace kazety
pRKL2-18	ORF6+7join_F1 ORF6+7join_R1	DNA viru <i>Vaccinia</i>	ORF6+7join_F1 ORF6+7join_R3	nesekvenováno	KanV2F in_ORF9_rev
	ORF6+7join_F2 ORF6+7join_R2	pGKL2		–	
	KL_orf6C_Flag2F ORF6+7join_R3	pRKL1-1			
pRKL2-19	inORF6_forv2 KL_orf6_split1R	pGKL2	inORF6_forv2 KL_orf6_split2R	KanV2F ORF7-K2_tail_2	KanV2F ORF7-K2_tail_2
	KL_orf6C_Flag2F KL_orf6_split2R	pRKL1-1		KanV2F	



**Obrázek 38:** Elektroforetogramy přípravy kmene *K. lactis* IFO1267\_pRKL2-18 a IFO1267\_pRKL2-19. Rekombinantní pGKL plasmidy jsou označeny červeně. M1: GeneRuler™ 1 kb DNA Ladder; M2: Lambda DNA/Eco130I (*StyI*) Marker, 16. (A) V případě kmene *K. lactis* IFO1267\_pRKL2-18 se transformací podařilo získat pouze kmen se zkráceným plasmidem pRKL2-18Δ, který podle další analýzy obsahoval s jistotou geny *G418<sup>R</sup>*, *K2ORF8* a *K2ORF9*.

(B) Kultivace dvou nezávislých monoklonií kmene *K. lactis* IFO1267\_pRKL2-19 v selektivních podmínkách vedla u všech analyzovaných klonů k ponechání plasmidu pGKL2. Plasmid pRKL2-19 tak není schopen v buňkách zastoupit funkci plasmidu pGKL2. Velikost pRKL2-18 14793 bp, velikost pRKL2-18Δ ~9000 bp, velikost pRKL2-19 14355 bp.

**Závěr:** Plasmid pRKL2-19 není schopen v buňkách zastoupit funkci plasmidu pGKL2. Výsledky naznačují, že fúze konzervovaných oblastí  $\beta$  a  $\beta'$  podjednotky RNAP v proteinu K2ORF6p je nezbytná pro udržení pGKL plasmidů v buňce. Plasmid pRKL2-18 zřejmě není technicky možné připravit *in vivo*.

### 5.1.9 Promotory pGKL plasmidů mají poxvirový původ

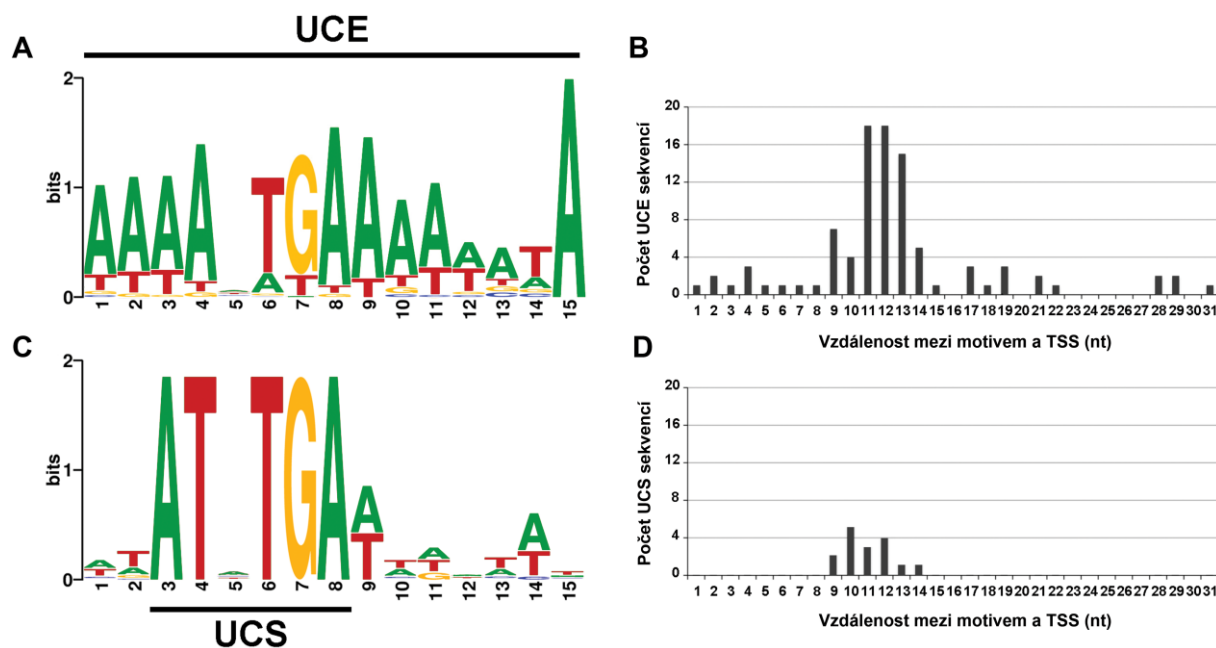
Ačkoliv byla sekvence promotorového elementu UCS (5'-ATNTGA-3'), esenciálního pro transkripci lineárních plasmidů, prvně identifikována již před mnoha lety (SOR a FUKUHARA 1985), nebyla popsána žádná sekvenční podobnost tohoto elementu se známými promotory RNA polymeráz, naznačující jeho fylogenetický původ. Jelikož je sekvence UCS poněkud krátká na hledání sekvenčně podobných elementů, použil jsem známé UCS předcházející všech 15 otevřených čtecích rámců pGKL plasmidů (STARK *et al.* 1990; LARSEN a MEINHARDT 2000) pro sestrojení manuálního sekvenčního porovnání včetně sekvencí sousedících s UCS po i proti směru transkripce o ~10 bp v programu BioEdit 7.0.5.3 a navrhl jsem nový konsenzus motiv pomocí výchozího nastavení programu WebLogo 2.8.2. Tento konsenzus jsem použil pro hledání podobných, literaturou popsaných, promotorových elementů, spojených s transkripcí vícepodjednotkových RNAP. Zvláště jsem se zaměřil na promotory virových RNAP, neboť moje fylogenetická analýza plasmidových RNAP naznačila virový původ těchto RNA polymeráz.

Pro hledání podobných promotorových elementů je kromě vlastní sekvence elementu velmi výhodná i informace o umístění elementu vzhledem k místu startu transkripce (TSS). Počáteční místa transkripce byla anotována pro několik ORF pGKL plasmidů na základě několika prací, kdy pro každý analyzovaný ORF bylo detekováno více po sobě jdoucích iniciačních nukleotidů transkripce (kapitola 2.1.1). V naší laboratoři však jsou díky experimentům Václava Vopálenského využívajícím rychlou amplifikaci 5' konců cDNA (5' RACE-PCR) dostupné informace o místech iniciace transkripce všech genů pGKL plasmidů. Dostupné sekvence (~25 sekvencí jednotlivých cDNA klonů pro každý ORF; 376 sekvencí celkem) jsem použil pro manuální anotaci TSS všech ORF pGKL plasmidů, kdy každému ORF jsem přiřadil pouze jediný (majoritní) TSS. Stručně řečeno, cDNA klony transkriptů genů *K2ORF2*, *K2ORF3* a *K2ORF8* vykazovaly vzor jediného vrcholového TSS, kde více než 60 % klonů mělo stejný iniciační nukleotid a každý jiný iniciátorový nukleotid nebyl zastoupen ve více než 25 % klonů. Sekvenování cDNA klonů odpovídajících 5' koncům transkriptů všech ostatních 12 otevřených čtecích rámců vykazovalo vzor vícečetných TSS v důsledku přítomnosti netemplátovaného 5' poly(A) leaderu heterogenní délky (1–21 nt). Pro tyto ORF byl jako TSS anotován první



adenosinový zbytek templátu, který v sekvenčním porovnání odpovídal prvnímu 5' adenosinovému zbytku cDNA klonu, jenž mohl vzniknout kontinuální transkripční templátu.

Podářilo se mi nalézt velkou sekvenční podobnost části prodloužené konzensus sekvence UCS motivu pGKL plasmidů (Obrázek 39C) s UCE sekvencí, což je promotorový element časných genů poxvirů (Obrázek 39A). Motiv UCE je 15 nt dlouhý AT bohatý element s jakýmkoli nukleotidem na 5. pozici, následovaným nukleotidy TGA (YANG *et al.* 2011a). Tyto vlastnosti motivu UCE rozšířený motiv UCS splňuje. Medián vzdálenosti 3' konce motivu UCE k anotovaným TSS 84 časných otevřených čtecích rámců viru *Vaccinia* je 12 nukleotidů (YANG *et al.* 2011a), viz Obrázek 39B. Medián vzdálenosti 3' konce rozšířeného UCS motivu k transkripčním startům 15 ORF pGKL plasmidů, anotovaných na základě 5' RACE-PCR experimentů, je 11 nukleotidů (Obrázek 39D). Kromě toho je TSS nukleotid všech genů pGKL plasmidů adenosinový zbytek, podobně jako v případě TSS časných genů poxvirů, kde jsou purinové báze dominantní nukleotidy TSS (YANG *et al.* 2011a).



**Obrázek 39:** Promotorový element UCS pGKL plasmidů vykazuje vysokou sekvenční podobnost s promotory časných genů poxvirů. **(A)** Konsenzus motiv UCE elementu promotorů časných genů viru *Vaccinia* vypočtený ze sekvencí 84 otevřených čtecích rámců, jenž byl identifikován na základě celogenomových analýz transkriptů VACV (YANG *et al.* 2011a). **(B)** Graf ukazující počet UCE sekvencí v závislosti na jejich vzdálenostech od příslušných míst startu transkripce 84 časných otevřených čtecích rámců VACV. Upraveno podle (YANG *et al.* 2011a). **(C)** Rozšířený konsenzus motiv promotorů 15 otevřených čtecích rámců pGKL plasmidů. Tento motiv obsahuje UCS element, který je konzervován mezi všemi kvasinkovými cytoplasmatickými lineárními plasmidy. **(D)** Graf ukazující počet rozšířených UCS sekvencí v závislosti na jejich vzdálenosti k příslušným TSS 15 otevřených čtecích rámců pGKL plasmidů. Místa startu transkripce byla anotována na základě 5' RACE-PCR experimentů naší laboratoře. Pro promotory pGKL genů, jejichž transkripty obsahovaly netemplátový 5' poly(A) leader, byl jako TSS považován první adenosinový zbytek templátu, který v sekvenčním porovnání odpovídal prvnímu 5' adenosinovému zbytku transkriptu, jenž mohl vzniknout kontinuálním přepisem templátu.

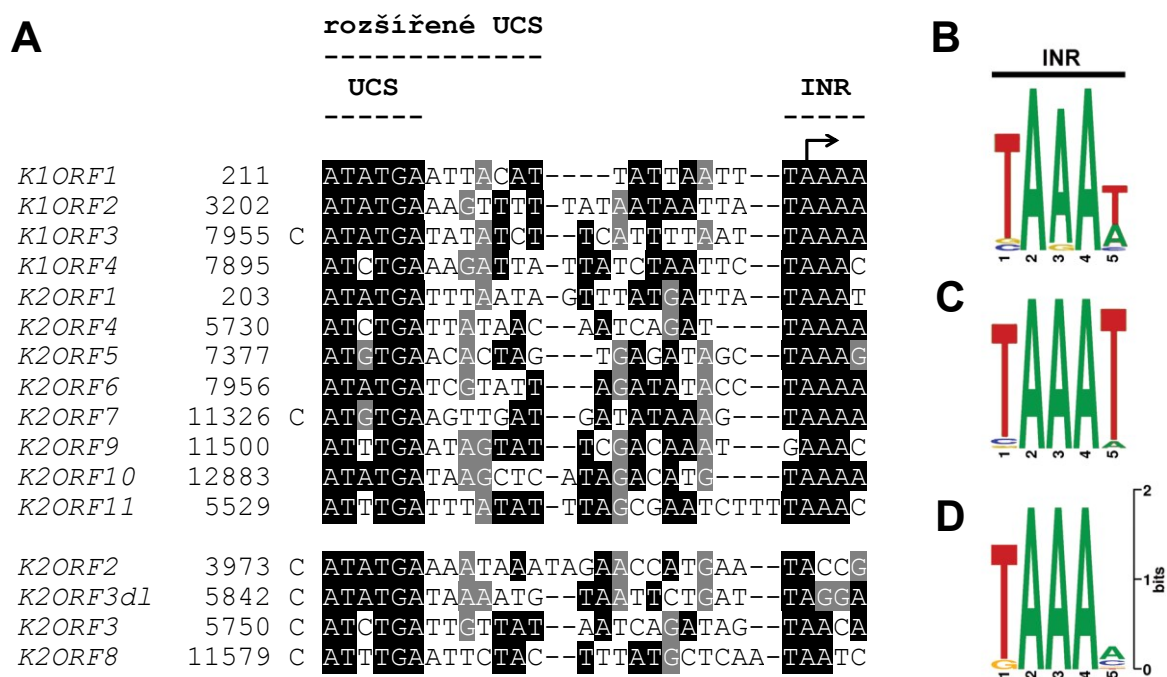


**Závěr:** Výsledky ukazují, že promotorové elementy časných genů poxvirů a genů lineárních plasmidů jsou podobné jak s ohledem na jejich sekvenci, tak i na jejich umístění vzhledem k místu iniciace transkripce, což poukazuje na pravděpodobně společný evoluční původ.

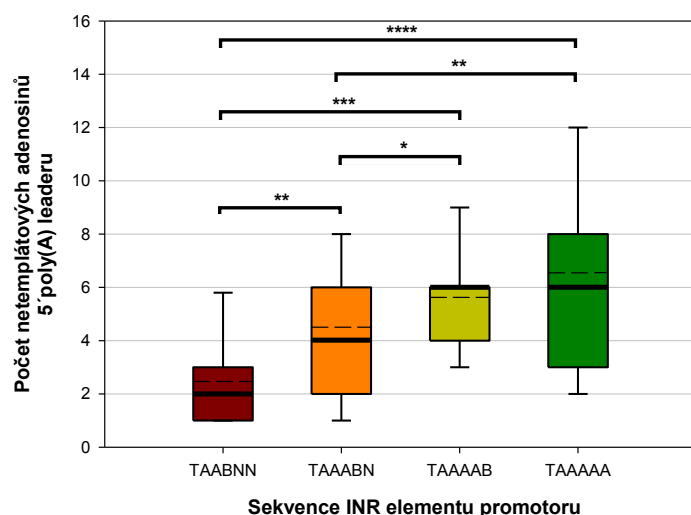
### 5.1.10 Promotory pGKL plasmidů obsahují iniciátorový region

Přítomnost netemplátového 5' poly(A) leaderu heterogenní délky (1–21 nt) u většiny cDNA klonů odpovídajícím 5' koncům transkriptů 12 genů pGKL plasmidů (detekovaná pomocí 5' RACE-PCR experimentů) vzbudila otázku, jakým mechanismem by mohl tento 5' poly(A) leader vznikat. Protože transkripty středních a pozdních genů VACV obsahují netemplátový 5' poly(A) leader, vznikající klouzáním RNAP po TAAAT iniciátorovém regionu promotoru *in vitro* a *in vivo* (kapitola 2.2.7), rozhodl jsem se hledat obdobný INR element v promotorech pGKL plasmidů. Pomocí bioinformatické analýzy se mi podařilo detekovat předpokládaný INR se sekvencí TAAAN u zmíněných 12 genů pGKL plasmidů (Obrázek 40B).

Na základě mutačních experimentů promotorů VACV se předpokládá, že pro efektivní syntézu 5' poly(A) leaderu mRNA klouzáním RNAP po templátu jsou potřeba alespoň tři následné adenosinové zbytky TAAAT elementu (DAVISON a MOSS 1989b). Promotory všech genů pGKL plasmidů, jejichž transkripty jsou většinou 5' polyadenylované, obsahují alespoň tři následné adenosinové zbytky v předpokládaném INR elementu (Obrázek 40A). Menší množství transkriptů genů *K2ORF3* (40 %) a *K2ORF8* (45 %) také obsahovalo netemplátový 5' poly(A) leader, přičemž v předpokládaném INR elementu promotorů těchto genů jsou dva následné adenosinové zbytky (Obrázek 40A). Předpokládaný INR element pGKL plasmidů má však oproti INR elementům viru *Vaccinia* značnou volnost nukleotidů na pozici 5, což znamená, že množství po sobě jdoucích adenosinových nukleotidů v templátové DNA promotoru může být běžně vyšší než tři. Při analýze sekvencí cDNA klonů transkriptů pGKL plasmidů získaných 5' RACE-PCR experimenty je také na první pohled patrný jev, že se vzrůstajícím množstvím adenosinových nukleotidů v INR elementu templátové DNA promotoru vzrůstá i počet netemplátově zařazených adenosinových nukleotidů 5' poly(A) leaderu mRNA (Obrázek 41 a Tabulka 22), což je v souladu s modelem klouzání RNAP po promotoru jakožto mechanismem zodpovědným za tvorbu 5' poly(A) leaderu mRNA pGKL plasmidů.



**Obrázek 40:** Promotory genů pGKL plasmidů obsahují předpokládaný iniciátorový region (INR). (A) Sekvenční porovnání promotorů pGKL plasmidů. Geny jsou rozděleny do dvou skupin v závislosti na přítomnosti (horní skupina) či nepřítomnosti (dolní skupina) netemplátového 5' poly(A) leaderu u většiny analyzovaných cDNA klonů získaných 5' RACE-PCR experimenty. *K2ORF3dl* značí dlouhou transkripční variantu genu *K2ORF3*, detekovanou pomocí 5' RACE-PCR. Předpokládané místo transkripčního startu, anotované na základě 5' RACE-PCR experimentů, je zobrazeno šipkou. Anotované místo iniciace transkripce u horní skupiny genů odpovídá prvnímu adenosinovému zbytku cDNA klonů s 5' poly(A) leaderem, jenž mohl vzniknout kontinuálním přepisem templátu. Sekvence odpovídající části rozšířeného UCS elementu a nově detekovanému iniciátorovému regionu byly porovnány ručně v programu BioEdit 7.0.5.3. Sekvence mezi těmito dvěma definovanými promotorovými motivy byly porovnány pomocí ClustalW 1.81 v prostředí BioEdit 7.0.5.3. Ve výsledném porovnání byla pomocí programu BoxShade 3.21 zvýrazněna sekvenční identita (černé stínování), pokud se stejný nukleotid vyskytoval u  $\geq 50\%$  sekvencí; sekvenční podobnost (šedé stínování), pokud se podobný nukleotid vyskytoval u  $\geq 50\%$  sekvencí. (B) Koncenzus motiv INR promotorů středních genů viru *Vaccinia* (YANG *et al.* 2012). (C) Koncenzus motiv INR promotorů pozdních genů viru *Vaccinia* (YANG *et al.* 2012). (D) Koncenzus motiv předpokládaného INR promotorů 12 genů pGKL plasmidů, jejichž cDNA klony většinou obsahovaly netemplátový 5' poly(A) leader. Koncenzus motiv byl vytvořen v programu WebLogo 2.8.2.



**Tabulka 22:** Počet netemplátových adenosinů 5' poly(A) leaderu mRNA pGKL plasmidů

INR	Počet sekvencí	Medián	Průměr
TAABNN	15	2	2,47
TAAABN	105	4	4,51
TAAAAB	67	6	5,63
TAAAAA	75	6	6,55

**Obrázek 41:** Vzrůstající počet netemplátových nukleotidů 5' poly(A) leaderu transkriptů pGKL plasmidů s počtem následných adenosinových zbytků v INR elementu promotoru. Krabicový graf počtu netemplátových adenosinů 5' poly(A) leaderu mRNA pGKL plasmidů v závislosti na sekvenci INR elementu. Spodní a vrchní část boxu indikuje první a třetí kvartil. Úsečky indikují 10. a 90. percentil. Odlehlé hodnoty nejsou vyznačeny. Medián je vyznačen tlustou čarou. Aritmetický průměr je vyznačen přerušovanou čarou. Normální rozdělení dat bylo testováno pomocí Shapiro-Wilk testu. Protože některé kategorie nesplňovaly normální rozdělení, byly data statisticky vyhodnoceny pomocí neparametrického Kruskal-Wallis testu následovaného *post-hoc* Dunn testem s Benjamini-Hochberg FDR korekcí. \*: hladina signifikance  $P < 0,05$ . \*\*: hladina signifikance  $P < 0,01$ . \*\*\*: hladina signifikance  $P < 0,001$ . \*\*\*\*: hladina signifikance  $P < 0,0001$ . B: jakýkoliv nukleotid kromě adenosinu; N: jakýkoliv nukleotid. Promotorovou sekvenci TAABNN (červeně) mají geny *K2ORF3* a *K2ORF8*. Promotorovou sekvenci TAAABN (oranžově) mají geny *K1ORF4*, *K2ORF1*, *K2ORF5*, *K2ORF9* (GAAABN) a *K2ORF11*. Promotorovou sekvenci TAAAAB (žlutě) mají geny *K1ORF2*, *K2ORF6* a *K2ORF10*. Promotorovou sekvenci TAAAAA (zeleně) mají geny *K1ORF1*, *K1ORF3*, *K2ORF4* a *K2ORF7*. Celkem bylo k analýze použito 262 sekvencí 5' RACE-PCR obsahujících netemplátové adenosinové nukleotidy. V ojedinělých případech některé sekvence v oblasti netemplátového 5' poly(A) leaderu obsahovaly jeden nukleotid jiný, než adenosin. Pro statistické účely bylo s jinými nukleotidy těchto sekvencí zacházeno tak, jako by se jednalo o adenosiny.

**Závěr:** Promotory genů pGKL plasmidů obsahují předpokládaný iniciátorový region, jenž může být zodpovědný za 5' polyadenylaci většiny plasmidových transkriptů *in vivo*.

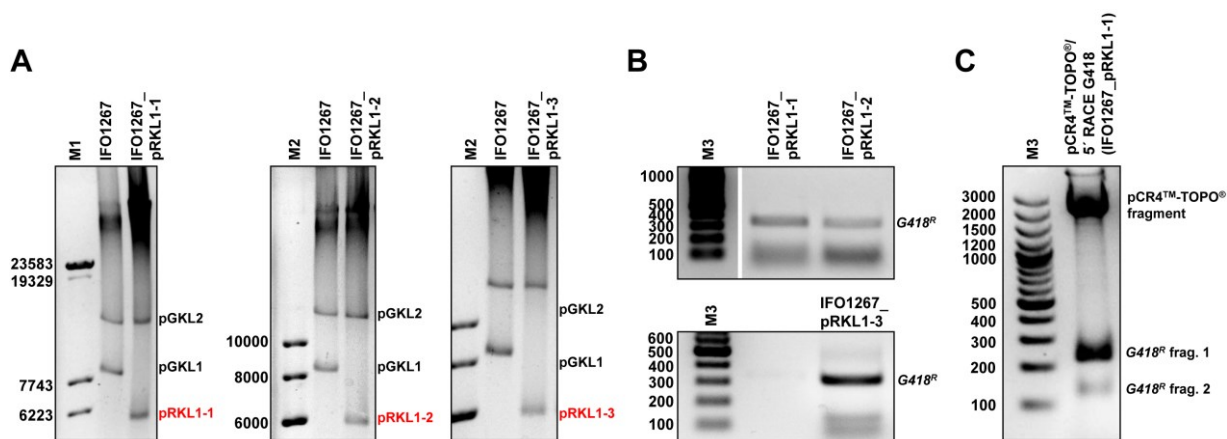
### 5.1.11 Klouzání RNA polymerázy v iniciátorovém regionu vede k 5' polyadenylaci plasmidových mRNA *in vivo*

Následně jsem se rozhodl otestovat, zda je předpokládaný INR element promotoru spojený s 5' polyadenylací transkriptů pGKL plasmidů. Byly připraveny tři kmeny *K. lactis* s modifikovanými pGKL plasmidy postupem již detailněji popsáním v kapitole 5.1.1.1 na příkladu plasmidu pRKL2-2. Připravené kmeny obsahovaly modifikovaný plasmid pGKL1 kódující gen rezistence ke G418 (vložený místo genu *K1ORF2*) pod kontrolou tří variant

promotoru K1UCR2, lišících se v sekvencích předpokládaných iniciátorových regionů: TAAAA (kmen IFO1267\_pRKL1-1), TAACA (kmen IFO1267\_pRKL1-2) a TACCA (kmen IFO1267\_pRKL1-3). Nejdůležitější informace k přípravě a charakterizaci nových kmenů souvisejících s touto kapitolou jsou shrnuty v Tabulce 23. Elektroforetogramy detekce DNA rekombinantních pGKL plasmidů v připravených kmenech jsou na Obrázku 42A.

Tabulka 23: Primery a templáty pro přípravu a charakterizaci kmenů této kapitoly					
Plasmid	Primery na přípravu kazety pomocí PCR	Templát pro přípravu kazety	Primery pro ověření integrace kazety a přípravu templátu na sekvenování / Sekvenace	Primery pro ověření integrace kazety 2	Primery pro ověření vyředění nemodifik. plasmidu
pRKL1-1	pGKL-kanF pGKL-kanR	pUG6	in_Kan_rev1 ORF1-K1_tail_2 in_Kan_rev1	KanVerF ORF3-pGKL1_tail	ORF1-K1_tail_2 5_RACE_O2_K1
pRKL1-2	K1O2_prom_del_F pGKL-kanR	pRKL1-1	in_Kan_rev1 ORF1-K1_tail_2 in_Kan_rev1	KanVerF ORF3-pGKL1_tail	ORF1-K1_tail_2 5_RACE_O2_K1
pRKL1-3	K1O2_prom_del2_F pGKL-kanR	pRKL1-1	in_Kan_rev1 ORF1-K1_tail_2 in_Kan_rev1	KanVerF ORF3-pGKL1_tail	ORF1-K1_tail_2 5_RACE_O2_K1

Z připravených kvasinkových kmenů byla izolována celková RNA (metoda 4.2.3.1.1). Zbývající DNA ve vzorcích byla odstraněna deoxyribonukleázou I (metoda 4.2.3.1.4) a kvalita RNA byla ověřena pomocí agarosové elektroforézy (metoda 4.2.3.3.1) a UV spektrofotometrie (4.2.3.2). Purifikovaná celková RNA byla použita k reverzní transkripci využívající náhodné primery (metody 4.2.3.6.1 a 4.2.3.6.2). Vzniklá cDNA byla purifikována (metoda 4.2.2.1.7) a použita k syntéze homopolymerního 3' konce (metoda 4.2.2.4.6). Vzniklá cDNA s 3' poly(G) koncem byla použita pro 5' RACE-PCR s primery in\_Kan\_rev1 a olig2(dC)anchor. Produkty 5' RACE-PCR reakcí byly rozděleny agarosovou elektroforézou (metoda 4.2.2.3.1), specifický produkt byl izolován (metoda 4.2.2.1.5) a klonován do vektoru pCR<sup>TM</sup>4-TOPO<sup>®</sup> (metoda 4.2.2.4.5). Vektory byly transformovány do buněk *E. coli* XL-1 Blue (metoda 4.2.2.5.1). Ze získaných monoklonií byly vektory následně izolovány (metoda 4.2.2.1.1) a charakterizovány pomocí restriční štěpení enzymy *NruI* a *EcoRI* (metoda 4.2.2.4.1) následovaného agarosovou elektroforézou (metoda 4.2.2.3.1). Z kolonií, jejichž vektory vykazovaly přítomnost insertu odpovídajícímu specifickému produktu 5' RACE-PCR reakce, byly vektory izolovány (metoda 4.2.2.1.2), jejich kvalita byla ověřena UV spektrofotometrií (4.2.2.2), a následně byly podrobeny sekvenaci DNA (metoda 4.2.2.8.4) v oblasti insertu primerem T3 nebo T7. Elektroforetogramy detekce produktů 5' RACE-PCR genu rezistence ke G418 (*G418<sup>R</sup>*) jsou na Obrázku 42B. Ukázkový elektroforetogram restriční analýzy vektoru pCR<sup>TM</sup>4-TOPO<sup>®</sup> s 5' RACE-PCR produktem genu rezistence ke G418 je na Obrázku 42C.

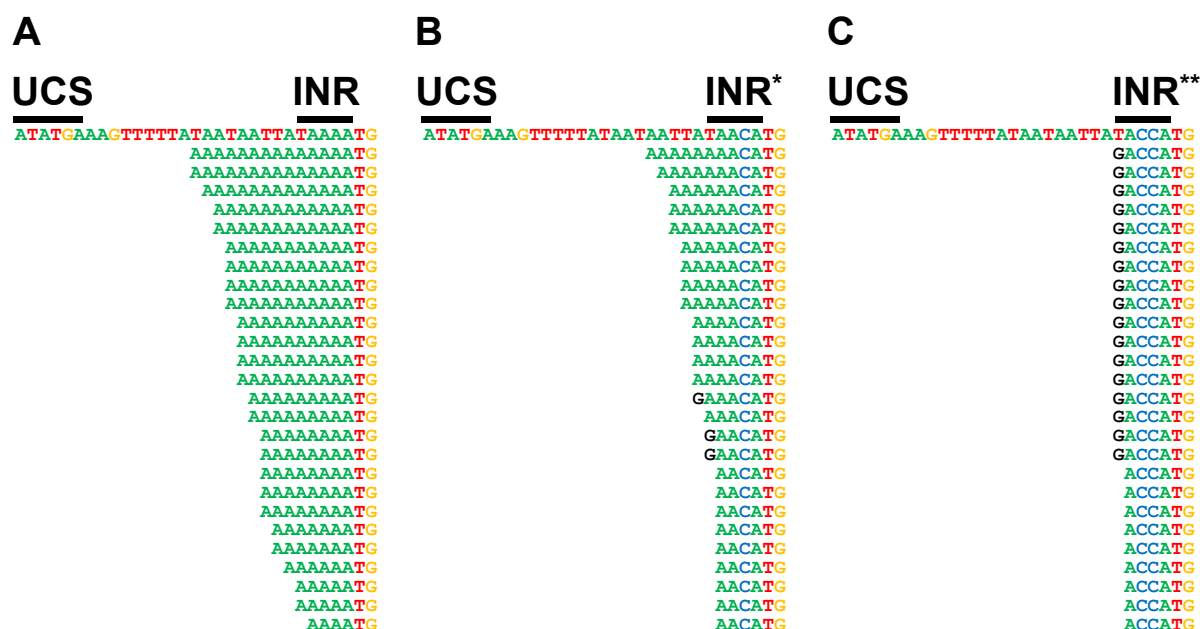


**Obrázek 42:** Příprava a analýza kmenů *K. lactis* za účelem zjištění vlivu INR elementu promotoru na netemplátovou 5' polyadenylaci mRNA pGKL plasmidů. (A) Elektroforetogramy DNA připravených kmenů kódujících gen rezistence ke G418 pod kontrolou různých variant promotoru K1UCR2. Rekombinantní pGKL plasmidy jsou označeny červeně. Velikost rekombinantních pGKL plasmidů je uvedena v Tabulce 6. Zobrazeny jsou vybrané monoklonie po pěti pasážích nezávislých klonů získaných transformací a následnou selekcí. M1: Lambda DNA/Eco130I (*StyI*) Marker, 16; M2: GeneRuler™ 1 kb DNA Ladder. (B) Elektroforetogramy 5' RACE-PCR genu rezistence ke G418 (*G418<sup>R</sup>*) z připravených kmenů *K. lactis* IFO1267\_pRKL1-1, IFO1267\_pRKL1-2 a IFO1267\_pRKL1-3. Specifický produkt (~329 bp) je označen vpravo. M3: GeneRuler™ 100 bp Plus DNA Ladder. (C) Elektroforetogram restrikční analýzy vektoru pCR4<sup>TM</sup>-TOPO<sup>®</sup> obsahujícího specifický produkt 5' RACE-PCR reakce genu *G418<sup>R</sup>* z celkové cDNA kmene *K. lactis* IFO1267\_pRKL1-1. Vektor byl štěpen enzymy *EcoRI* a *NruI*. Velikost fragmentu odpovídajícímu vektoru je 3938 bp, velikost fragmentů odpovídajícím štěpenému produktu 5' RACE-PCR genu *G418<sup>R</sup>* je v závislosti na orientaci insertu buď ~217 bp a 130 bp nebo 221 bp a ~126 bp. Přibližně 20 vektorů vykazujících stejný restrikční vzor bylo sekvenováno pro každý z uvedených kvasinkových kmenů.

Výsledky ukázaly, že 5' poly(A) leader transkriptů genu rezistence ke G418 byl přítomen, když sekvence K1UCR2 obsahovala předpokládaný INR divokého typu (TAAAA) u kmene *K. lactis* IFO1267\_pRKL1-1 (Obrázek 43A). Při použití mutovaného TAACA iniciátorového regionu (kmen IFO1267\_pRKL1-2) promotoru K1UCR2 byla délka netemplátového 5' poly(A) leaderu *G418<sup>R</sup>* transkriptů značně redukována (Obrázek 43B). Když byl použit promotor K1UCR2 s dvěma bodovými mutacemi v INR (TACCA; kmen IFO1267\_pRKL1-3), netemplátový 5' poly(A) leader mRNA genu *G418<sup>R</sup>* nebyl přítomen (Obrázek 43C).

Pro přípravu templátové cDNA z celkové RNA kvasinek byla použita reverzní transkriptáza SuperScript III, která patří do skupiny reverzních transkriptáz, které jsou schopny překonat 5'–5' vazbu mezi mRNA čepičkou a prvním nukleotidem mRNA, což umožňuje detekovat přítomnost mRNA čepičky transkriptu na úrovni cDNA (SCHMIDT a MUELLER 1999). V kvasinkových kmenech s mutovanými INR elementy se znatelně zvýšilo množství transkriptů *G418<sup>R</sup>* s mRNA čepičkou (Obrázek 43B a C). To značí, že část detekovaných transkriptů z kmene *K. lactis* IFO1267\_pRKL1-2 a většina detekovaných transkriptů kmene *K. lactis*

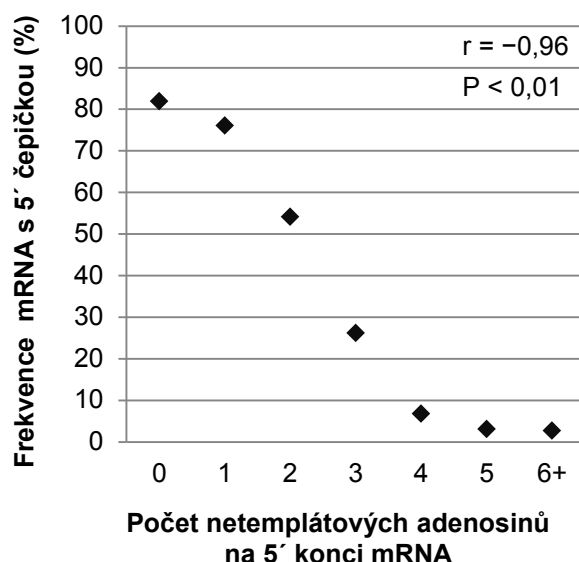
IFO1267\_pRKL1-3 pravděpodobně odpovídá nezkráceným transkriptům, tudíž nepřítomnost netemplátového 5' poly(A) leaderu těchto mRNA nemůže být vysvětlena 5' nukleolytickou degradací RNA. Tyto výsledky jsou také ve shodě s již dříve zaznamenanou signifikantní negativní korelací mezi délkou 5' poly(A) leaderu a přítomností 5' čepičky na transkriptech plasmidových genů (Tabulka 24 a Obrázek 44).



**Obrázek 43:** Promotory pGKL plasmidů obsahují iniciátorový region (INR) odpovídající za netemplátovou 5' polyadenylaci mRNA. (A) 5' RACE-PCR analýza genu *G418<sup>R</sup>* z kmene *K. lactis* IFO1267\_pRKL1-1. V tomto a následujících panelech odpovídá horní sekvence templátové (plasmidové) DNA s vyznačenou sekvencí UCS a INR; sekvence umístěné níže představují jednotlivé sekvenované cDNA klony (5' nepřekládaná oblast je zobrazena v plném rozsahu až do start kodónu translace, ATG). Guanosinové zbytky odpovídající původním 5' mRNA čepičkám, přítomné v některých klonech cDNA, jsou vyznačeny černě. (B) 5' RACE-PCR analýza genu *G418<sup>R</sup>* z kmene *K. lactis* IFO1267\_pRKL1-2 nesoucího jednobodovou mutaci (\*) v předpokládaném INR promotoru, snižující počet po sobě jdoucích adenosinových zbytků v templátu. (C) 5' RACE-PCR analýza genu *G418<sup>R</sup>* z kmene *K. lactis* IFO1267\_pRKL1-3 nesoucího dvoubodovou mutaci (\*\*) v předpokládaném INR promotoru, která ruší následné adenosinové zbytky v templátu.

Tabulka 24: Četnost 5' mRNA čepiček u transkriptů pGKL plasmidů s různým počtem netemplátových adenosinů 5' poly(A) leaderu			
Počet netemplátových adenosinů	Počet sekvencí	Počet sekvencí s 5' čepičkou	Frekvence výskytu 5' čepičky
0	88	72	82 %
1	25	19	76 %
2	37	20	54 %
3	23	6	26 %
4	30	2	7 %
5	33	1	3 %
6+	114	3	2,6 %





**Obrázek 44:** Signifikantní negativní korelace mezi počtem netemplátových adenosinů 5' poly(A) leaderu a přítomností 5' mRNA čepičky na transkriptech pGKL plasmidů. Bodový graf znázorňuje frekvenci výskytu 5' mRNA čepičkových struktur u transkriptů jednotlivých tříd. Písmeno *r* značí hodnotu Pearsonova korelačního koeficientu. Výsledky jsou signifikantní na hladině  $P < 0,01$ . Celkem bylo k analýze použito 350 sekvencí cDNA klonů získaných pomocí 5' RACE-PCR. Přibližně 7 % sekvencí cDNA klonů bylo z analýzy vyřazeno, protože neobsahovaly anotovaný nukleotid místa transkripčního startu. Tyto sekvence byly vyřazeny kvůli podezření, že se mohlo jednat o částečně 5' degradované transkripty, jenž by postrádaly

mRNA čepičku. Protože detekovaný počet transkriptů s šesti a více netemplátovými adenosiny 5' poly(A) leaderu obsahujících mRNA čepičku byl velmi nízký, byly tyto transkripty s dlouhým 5' poly(A) leaderem sloučeny do jedné kategorie. V ojedinělých případech některé sekvence v oblasti netemplátového 5' poly(A) leaderu obsahovaly jeden nukleotid jiný, než adenosin. Pro statistické účely bylo s jinými nukleotidy těchto sekvencí zacházeno tak, jako by se jednalo o adenosiny.

**Závěr:** Výsledky jasně ukazují, že iniciátorový region promotoru je spojený s 5' polyadenylací transkriptů pGKL plasmidů *in vivo*. Stejně jako obdobné výsledky získané u viru *Vaccinia in vivo* (DAVISON a MOSS 1989b), naše výsledky nasvědčují, že přidávání netemplátových adenosinů 5' poly(A) leaderu mRNA pGKL plasmidů probíhá mechanismem klouzání RNA polymerázy v INR elementu promotoru. Sekvence iniciátorového regionu promotoru tvoří nezávislý DNA prvek, který není ovlivněn sekvencí genu, protože vzor 5' polyadenylace sekvenovaných cDNA klonů transkriptů genů *K1ORF2* a *G418<sup>R</sup>* vznikajících z K1UCR2 promotoru s INR divokého typu byl stejný (data nezobrazena). Výsledky z nově připravených rekombinantních kmenů s mutovanými promotory také opět poukázaly na signifikantní negativní korelaci mezi počtem netemplátových adenosinů 5' poly(A) leaderu a přítomností 5' mRNA čepičky na transkriptech pGKL plasmidů, která byla detekována již dříve na transkriptech vznikajících z promotorů divokých typů. Tento fenomén tak s největší pravděpodobností není ovlivněn sekvencí genu.

### 5.1.12 Transkripty pGKL plasmidů obsahují 5' čepičkovou strukturu

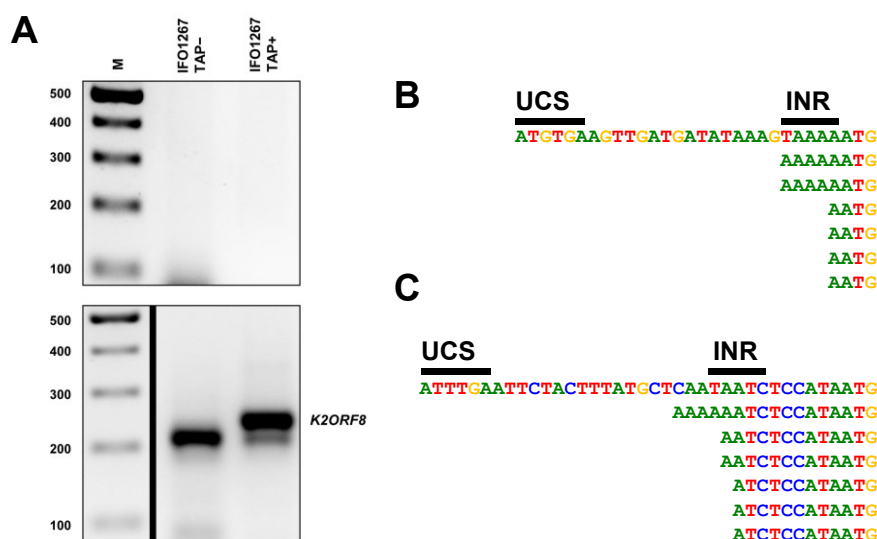
Přítomnost 5' mRNA čepičky na transkriptech pGKL plasmidů byla dlouhou dobu jen předpokládána díky tomu, že gen *K2ORF3* kóduje mRNA *capping* enzym, který je svým složením domén podobný mRNA *capping* enzymu viru *Vaccinia* (LARSEN *et al.* 1998; TIGGEMANN *et al.*

2001). Výsledky mého školitele využívající enzymy hDcp2 nebo Rai1 odstraňující 5' čepičku z mRNA a následnou 5' RACE-PCR analýzu při použití reverzní transkriptázy SuperScript III naznačují, že alespoň část transkriptů všech genů pGKL plasmidů obsahuje klasickou N<sup>7</sup>-methylguanosinovou čepičkovou strukturu typu *cap* 0. Nezávisle na těchto experimentech jsem se rozhodl ověřit přítomnost 5' čepičkové struktury transkriptů pGKL plasmidů pomocí 5' RACE-PCR zprostředkované RNA ligázou (5' RLM-RACE). K tomuto účelu byla použita souprava FirstChoice<sup>®</sup> RLM-RACE Kit (Invitrogen), jejíž chemikálie mi byly laskavě poskytnuty Liborem Krásným. Metoda RLM-RACE (též známá jako *oligocapping*) využívá sled tří následných enzymatických reakcí, díky kterým dojde k nahrazení 5' čepičkové struktury mRNA krátkým RNA oligonukleotidem před reverzní transkripcí, a následné detekci odpovídající cDNA pomocí PCR (MARUYAMA a SUGANO 1994). K odstranění 5' čepičkové struktury mRNA se využívá kyselá pyrofosfatáza tabáku, která je schopna štěpit 5'-5' vazbu a odstranit tak 5' blokuující strukturu mRNA, ačkoliv zřejmě není specifická pro N<sup>7</sup>-methylguanosin (SHINSHI *et al.* 1976a; SHINSHI *et al.* 1976b). Touto metodou jsem se rozhodl detekovat transkripty genů *K2ORF7* a *K2ORF8*, přičemž transkripty genu *K2ORF7* z analýz V. Vopálenského patřily mezi transkripty s nejdelším 5' poly(A) leaderem (medián netemplátových adenosinů = 7) a nízkou četností mRNA čepičky (4,3 %), a transkripty genu *K2ORF8* z analýz V. Vopálenského patřily mezi transkripty s velmi krátkým 5' poly(A) leaderem (medián netemplátových adenosinů = 0) a vysokou četností mRNA čepičky (86 %).

Z kmene *K. lactis* IFO1267 byla izolována celková RNA (metoda 4.2.3.1.1). Zbývající DNA ve vzorcích byla odstraněna deoxyribonukleázou I (metoda 4.2.3.1.4) a kvalita RNA byla ověřena pomocí agarosové elektroforézy (metoda 4.2.3.3.1) a UV spektrofotometrie (4.2.3.2). Purifikovaná celková RNA byla použita k RLM-RACE experimentům (4.2.3.4.1) a následně k reverzní transkripci využívající náhodné primery (metody 4.2.3.6.1 a 4.2.3.6.2). Vzniklá cDNA byla purifikována (metoda 4.2.2.1.7) a použita pro 5' RACE-PCR s primery uvedenými v Tabulce 25. Produkty 5' RACE-PCR reakcí byly rozděleny agarosovou elektroforézou (metoda 4.2.2.3.1). Protože specifické produkty nebyly patrné, byly produkty 5' RACE-PCR reakcí použity jako templáty pro druhé kolo PCR s primery uvedenými v Tabulce 25. Produkty 5' RACE-PCR reakcí byly rozděleny agarosovou elektroforézou a specifické produkty byly izolovány (metoda 4.2.2.1.5) a klonovány do vektorů pCR<sup>™</sup>4-TOPO<sup>®</sup> (metoda 4.2.2.4.5) a pJET1.2/blunt (metoda 4.2.2.4.4). Vektory byly transformovány do buněk *E. coli* XL-1 Blue (metoda 4.2.2.5.1). Ze získaných monoklonií byly vektory následně izolovány (metoda 4.2.2.1.1) a charakterizovány pomocí restriční štěpení (metoda 4.2.2.4.1) následovaného agarosovou elektroforézou (metoda 4.2.2.3.1). Z kolonií, jejichž vektory vykazovaly přítomnost insertu odpovídajícímu specifickému produktu 5' RACE-PCR reakce, byly vektory izolovány (metoda

4.2.2.1.2), jejich kvalita byla ověřena UV spektrofotometrií (4.2.2.2), a následně byly podrobeny sekvenaci DNA (metoda 4.2.2.8.5) v oblasti insertu primery uvedenými v Tabulce 25. Ukázkový elektroforetogram detekce produktu 5' RACE-PCR genu *K2ORF8* po prvním a druhém kole PCR je na Obrázku 45A. Pomocí této metody se podařilo detekovat specifické transkripty genů *K2ORF7* a *K2ORF8*, které prokazatelně obsahovaly 5' čepičkovou strukturu, protože tyto transkripty nebyly detekovány v kontrolní reakci bez kyselé pyrofosfatázy tabáku (Obrázek 45).

Tabulka 25: 5' RLM-RACE analýza transkriptů pGKL plasmidů						
Gen	Primery pro první kolo 5' RACE-PCR (35 cyklů)	Primery pro druhé kolo 5' RACE-PCR (35 cyklů)	Produkt	Klonováno do vektoru	Restrikční analýza vektoru	Sekvenace vektoru
<i>K2ORF7</i>	5' RACE Outer Primer 5RACE_O7_K2	5' RACE Inner Primer 5RACE_O7_K2	316 bp	pCR™ 4-TOPO®	<i>EcoRI</i> <i>XmiI</i>	T7
<i>K2ORF8</i>	5' RACE Outer Primer 5RACE_O8_K2	5' RACE Inner Primer 5RACE_O8_K2	246 bp	pCR™ 4-TOPO®	<i>EcoRI</i>	T7
				pJET1.2/ blunt	<i>EcoRI</i> <i>BglIII</i>	T7 5RACE_O8_K2



**Obrázek 45:** Transkripty pGKL plasmidů obsahují 5' čepičkovou strukturu. (A) Elektroforetogram 5' RLM-RACE analýzy transkriptů genu *K2ORF8* z kmene *K. lactis* IFO1267. Horní panel zobrazuje elektroforetickou analýzu produktů prvního kola 5' RACE-PCR, dolní panel zobrazuje elektroforetickou analýzu produktů druhého kola 5' RACE-PCR. Reakce, kdy k RNA byla (TAP+) a nebyla (TAP-) přidána kyselá pyrofosfatáza tabáku jsou označeny. Specifický produkt (246 bp) je označen vpravo. M: GeneRuler™ 100 bp Plus DNA Ladder. (B) 5' RLM-RACE analýza genu *K2ORF7* z kmene *K. lactis* IFO1267. V tomto a následujících panelech odpovídá horní sekvence templátové (plasmidové) DNA s vyznačenou sekvencí UCS a INR; sekvence umístěné níže představují jednotlivé sekvenované cDNA klony (5' nepřekládaná oblast je zobrazena v plném rozsahu až do start kodónu translace, ATG). (C) 5' RLM-RACE analýza genu *K2ORF8* z kmene *K. lactis* IFO1267.

**Závěr:** Nezávisle na výsledcích mého školitele se mi podařilo prokázat přítomnost 5' čepičkové struktury na transkriptech dvou vybraných genů pGKL plasmidů s využitím kyselé pyrofosfatázy tabáku. Ačkoliv je množství analyzovaných klonů malé, výsledky jsou ve shodě s dříve

detekovanou signifikantní negativní korelaci mezi počtem netemplátových adenosinů 5' poly(A) leaderu a přítomností 5' mRNA čepičky na transkriptech pGKL plasmidů. U transkriptů genu *K2ORF7*, které mají obvykle dlouhý netemplátový 5' poly(A) leader, byly metodou RLM-RACE detekovány transkripty obsahující 5' mRNA čepičku a velmi krátký či žádný netemplátový 5' poly(A) leader.

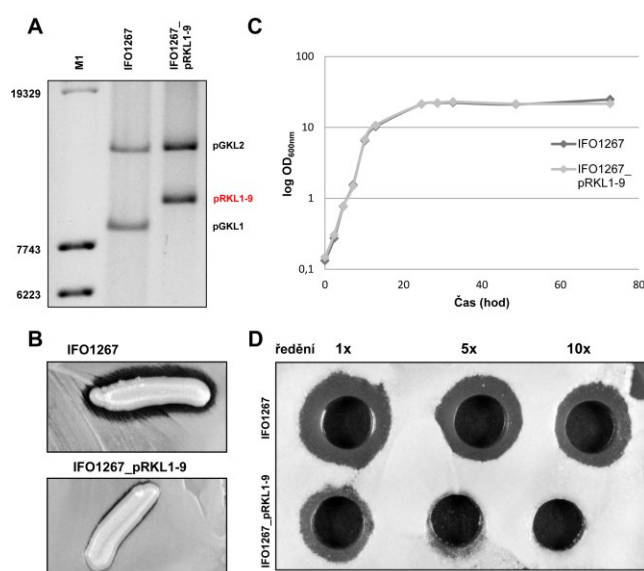
### 5.1.13 Mutace v iniciátorovém regionu promotoru vede ke snížené expresi genu *in vivo*

Dlouhodobým cílem naší laboratoře je popsat mechanismus iniciace translace mRNA pGKL plasmidů. Dosavadní výsledky ukázaly častou přítomnost 5' poly(A) leaderu a zároveň sporadickou přítomnost 5' mRNA čepičky na transkriptech pGKL plasmidů, přičemž se zdá, že translace plasmidových mRNA kódujících *killer* toxin je nezávislá na eukaryotickém translačním iniciačním faktoru 4E (eIF4E), což je faktor interagující s 5' mRNA čepičkou (ALTMANN *et al.* 1987; VOPÁLENSKÝ *et al.* 2018). Bylo popsáno, že 5' poly(A) leadery mohou za určitých podmínek u buněčných a virových mRNA fungovat jako eukaryotické translační enhancery *in vitro* a *in vivo* (GUDKOV *et al.* 2005; GILBERT *et al.* 2007; SHIROKIKH a SPIRIN 2008; XIA *et al.* 2011; DHUNGEL *et al.* 2017). Rozhodl jsem se otestovat, zda mutace promotoru, rušící následné adenosinové zbytky v INR elementu, u genu *K1ORF4* kódujícího  $\gamma$  podjednotku toxinu ovlivní expresi *killer* toxinu. Byl připraven kmen *K. lactis* s modifikovaný plasmidem pGKL1 postupem detailněji popsaným v kapitole 5.1.1.1. Připravený kmen IFO1267\_pRKL1-9 obsahoval modifikovaný plasmid pGKL1 kódující gen *K1ORF4* pod kontrolou promotoru K1UCR4 s TACCC iniciátorovým regionem. Nejdůležitější informace k přípravě a charakterizaci tohoto kmene jsou shrnuty v Tabulce 26. Elektroforetogram detekce DNA rekombinantního pGKL plasmidu je na Obrázku 46A. S přípravou tohoto kmene asistovala má bývalá studentka Jitka Vojáčková a na charakterizaci fenotypu tohoto kmene se podílel Václav Vopálenský.

Tabulka 26: Primery a templáty pro přípravu a charakterizaci kmene této kapitoly					
Plasmid	Primery na přípravu kazety pomocí PCR	Templát pro přípravu kazety	Primery na přípravu kazety pomocí fúzní PCR	Primery pro přípravu templátu na sekvenování / Sekvenace	Primery pro ověření integrace kazety
pRKL1-9	5RACE_O3_K1_2 K1_UCRmut-O4_R1	pGKL1	5RACE_O3_K1_2 K1_UCRmut-O4_R2	5_RACE_O4_K1 Hygro_tail	5_RACE_O4_K1 Hygro_tail
	K1_HA-ORF3_F2 K1_UCRmut-O4_R2	pcDNA™ 5/FRT/ TO		Hygro_tail	

Pro jednoduché testování *killer* fenotypu buněk kmenů *K. lactis* IFO1267 a IFO1267\_pRKL1-9 byly buňky těchto kmenů nanášeny na YPDA misku obsahující

$1 \times 10^6$  buněk podkladového kvasinkového kmene *S. cerevisiae* S6/1 citlivého na působení pGKL toxinu a inkubovány přes noc při 28 °C. Buňky obou kmenů *K. lactis* vykazují *killer* fenotyp, ačkoliv inhibiční zóna u kmene IFO1267\_pRKL1-9 je znatelně redukována (Obrázek 46B). Pro přesnější testování *killer* fenotypu byly oba kmeny kultivovány v YPD mediu při teplotě 28 °C po dobu přibližně 72 hodin a jejich optická denzita při 600 nm a produkce toxinu byla v pravidelných intervalech sledována. Růstové vlastnosti obou kmenů byly srovnatelné (Obrázek 46C). Produkce *killer* toxinu do kultivačního média byla testována jamkovým testem na YPDA agarových miskách s podkladem  $1 \times 10^6$  buněk citlivého kmene *S. cerevisiae* S6/1. Médium po kultivaci bylo zfiltrováno 0,2  $\mu$ m filtrem a 100  $\mu$ l média v různých ředěních bylo nanášeno do jamek YPDA misek, které byly inkubovány přes noc při 24 °C. Produkce *killer* toxinu kmene *K. lactis* IFO1267 byla po celou dobu kultivace ve všech odebraných vzorcích vyšší, než produkce *killer* toxinu kmene *K. lactis* IFO1267\_pRKL1-9 (Obrázek 46D).



**Obrázek 46:** Mutace v iniciátorovém regionu promotoru genu *K1ORF4* vede k nižší produkci *killer* toxinu pGKL plasmidů. **(A)** Elektroforetogram DNA připraveného kmene. Rekombinantní pGKL plasmid je označen červeně. Velikost plasmidu pRKL1-9 je 9988 bp. Zobrazena je vybraná monoklonie po pěti pasážích jednoho z klonů získaných transformací a následnou selekcí. M1: Lambda DNA/Eco130I (*StyI*) Marker, 16. **(B)** Test *killer* fenotypu buněk kmene *K. lactis* IFO1267 a IFO1267\_pRKL1-9 na podkladu buněk kmene *S. cerevisiae* S6/1 na agarové půdě YPDA. Inhibiční zóna kolem biomasy buněk je vidět černě. **(C)** Semilogaritmické zobrazení růstových

křivek kmenů *K. lactis* IFO1267 a IFO1267\_pRKL1-9 v YPD mediu při 28 °C. Jednotlivé body měření OD<sub>600nm</sub> a odběru média jsou zobrazeny. **(D)** Produkce *killer* toxinu v mediu po kultivaci buněk kmene *K. lactis* IFO1267 a IFO1267\_pRKL1-9 detekovaná jamkovým testem na podkladu buněk kmene *S. cerevisiae* S6/1 na agarové půdě YPDA. Ukázán je pouze vzorek 100  $\mu$ l média v různých ředěních, odebraný po 10 hodinách kultivace (OD<sub>600nm</sub> = 8). Ostatní vzorky se chovaly obdobně. Inhibiční zóna kolem jamky agarové půdy je vidět černě.

**Závěr:** Mutace v iniciátorovém regionu snižující počet následných adenosinových nukleotidů v promotoru genu *K1ORF4* kódujícím  $\gamma$  podjednotku toxinu vede ke snížení produkce *killer* toxinu. Obdobné mutace INR u postreplikativních promotorů viru *Vaccinia* vedou k inhibici exprese reportérového genu (DAVISON a MOSS 1989b; BALDICK *et al.* 1992; KNUTSON *et al.* 2006). Za předpokladu stejné hladiny syntézy mRNA z promotoru K1UCR4 divokého typu a promotoru K1UCR4 s mutovaným iniciátorovým regionem lze rozdíly v produkci toxinu



vysvětlit rozdílnou mírou translace 5' polyadenylovaných a 5' nepolyadenylovaných transkriptů genu *K1ORF4*.

#### 5.1.14 Časný promotor viru *Vaccinia* s iniciátorovým regionem nefunguje jako promotor pGKL plasmidů *in vivo*

Zaznamenaná sekvenční podobnost promotorového elementu UCE viru *Vaccinia* a elementu UCS pGKL plasmidů (kapitola 5.1.9) mě přiměla k otázce, zda by některý z promotorů časných genů viru *Vaccinia* mohl fungovat jako funkční promotor pro expresi reportérového genu z pGKL plasmidů. Většina promotorů pGKL plasmidů však obsahuje iniciátorový region, který zpravidla není součástí promotorů časných genů viru *Vaccinia*. Pomocí celogenomových analýz transkriptů VACV však byly nalezeny promotory 18 časných genů viru *Vaccinia*, které iniciátorový region obsahují a dávají vzniknout 5' polyadenylovaným transkriptům *in vivo* (YANG *et al.* 2011a). Rozhodl jsem se připravit tři kmeny *K. lactis* s modifikovaným plasmidem pGKL2 postupem detailněji popsáním v kapitole 5.1.1.1. Připravované kmeny měly obsahovat modifikovaný plasmid pGKL2 kódující gen *G418<sup>R</sup>* pod kontrolou promotoru genu *A5R*, *D12L* (obsahují INR) a *I4L* (bez INR) viru *Vaccinia*. Rezistence ke G148 byla zároveň použita k selekci buněk po transformaci. Zmíněným postupem se mi navzdory četným pokusům nikdy nepodařilo po transformaci získat klony pro další charakterizaci (data nezobrazena).

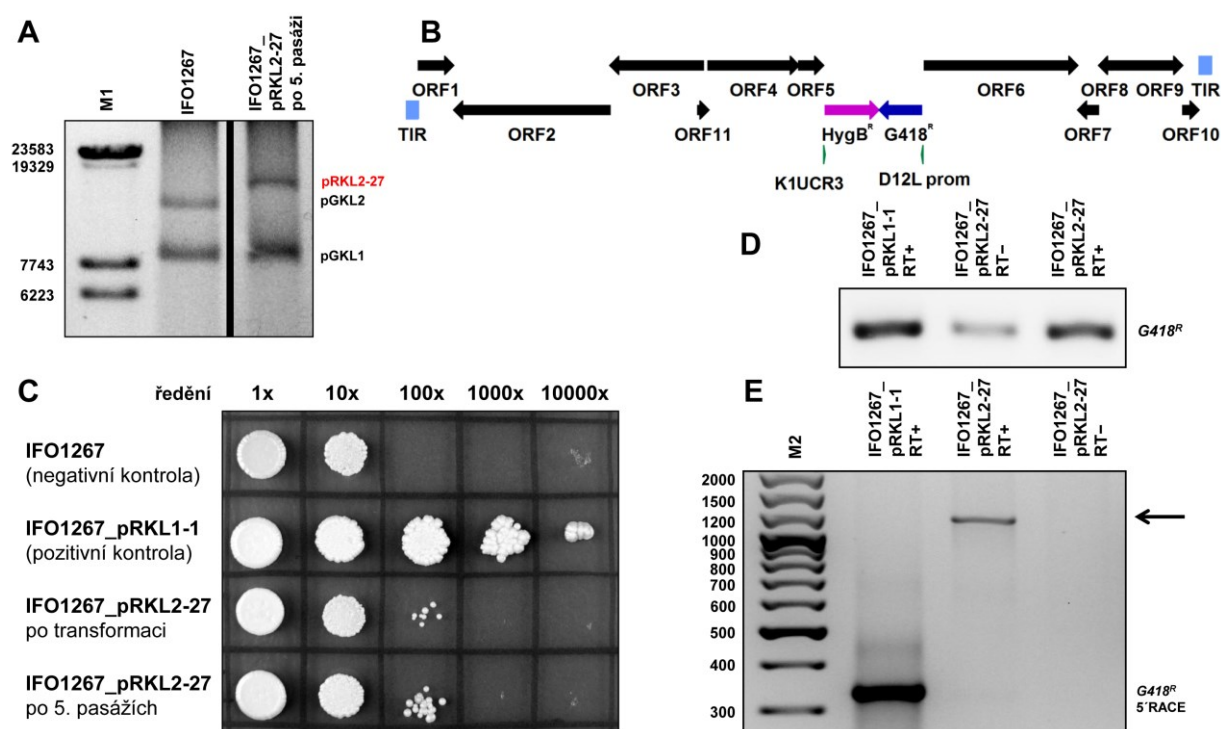
V dalším experimentu jsem se rozhodl připravit dva kmeny *K. lactis* s modifikovaným plasmidem pGKL2, které by umožňovaly detekci nízké hladiny exprese reportérového genu z promotorů VACV umístěných v pGKL plasmidech. Návrh transformační kazety spočíval ve vnesení genu rezistence k hygromycinu B (*HygB<sup>R</sup>*) pod kontrolou plasmidového promotoru společně s genem *G418<sup>R</sup>* pod kontrolou vakciniového promotoru *D12L* či *I4L*. K selekci klonů po transformaci pak sloužila rezistence k hygromycinu B. Tímto způsobem se podařilo připravit kmen *K. lactis* IFO1267\_pRKL2-27. Nejdůležitější informace k přípravě a charakterizaci tohoto kmene jsou shrnuty v Tabulce 27. Elektroforetogram detekce rekombinantního pGKL plasmidu a schéma jeho kódující kapacity je na Obrázku 47A a B. Na přípravě a částečně i na fenotypové charakterizaci tohoto kmene se podílela Jitka Vojáčková.

**Tabulka 27:** Primery a templáty pro přípravu a charakterizaci kmene této kapitoly

Plasmid	Primery na přípravu kazety pomocí PCR	Templát pro přípravu kazety	Primery na přípravu kazety pomocí fúzní PCR	Primery pro přípravu templátu na sekvenování / Sekvence	Primery pro ověření integrace kazety
pRKL2-27	pGKL_HA-ORF6_F1 pGKL_VV_prom3_R	pcDNA™ 5/FRT/TO	pGKL_HA-ORF6_F1 pGKL_D12L_prom_R	in_Kan_rev1 5RACE_O6_K2	in_Kan_rev1 5RACE_O6_K2
	kanR2 pGKL_D12L_prom_R	pGKL1		in_Kan_rev1	



Buňky připraveného kmene *K. lactis* IFO1267\_pRKL2-27 byly testovány na rezistenci ke G418 pomocí kapkového testu. V tomto testu vykazovaly buňky připraveného kmene mírně lepší růstové vlastnosti než buňky kmene IFO1267, které byly patrné zejména při použití nízkých koncentrací G418 a agarových půd se syntetickým médiem SD– (Obrázek 47C). Z připravených kvasinkových kmenů byla izolována celková RNA (metoda 4.2.3.1.1). Zbývající DNA ve vzorcích byla odstraněna deoxyribonukleázou I (metoda 4.2.3.1.4) a kvalita RNA byla ověřena pomocí agarosové elektroforézy (metoda 4.2.3.3.1) a UV spektrofotometrie (4.2.3.2). Purifikovaná celková RNA byla použita k reverzní transkripci využívající náhodné primery (metody 4.2.3.6.1 a 4.2.3.6.2). Vzniklá cDNA byla purifikována (metoda 4.2.2.1.7) a použita k detekci cDNA genu *G418<sup>R</sup>* pomocí PCR (Obrázek 47D), a také k syntéze homopolymerního 3' konce (metoda 4.2.2.4.6). Vzniklá cDNA s 3' poly(G) koncem byla použita pro 5' RACE-PCR analýzu cDNA genu *G418<sup>R</sup>* s primery in\_Kan\_rev1 a olig2(dC)anchor. Elektroforetogram detekce produktu 5' RACE-PCR genu *G418<sup>R</sup>* je na Obrázku 47E.



**Obrázek 47:** Časný promotor viru *Vaccinia* nefunguje jako promotor v pGKL plasmidech. (A) Elektroforetogram DNA připraveného kmene. Rekombinantní pGKL plasmid je označen červeně. Velikost plasmidu pRKL2-27 je 15359 bp. M1: Lambda DNA/Eco130I (*StyI*) Marker, 16. (B) Schématické znázornění kódující kapacity plasmidu pRKL2-27. Šipky indikují otevřené čtecí rámce a jejich transkripční směr. Tyrkysové obdélníky indikují terminální invertované repetice (TIR). Fialová šipka indikuje gen *HygB<sup>R</sup>* a modrá šipka indikuje gen *G418<sup>R</sup>*. Promotory K1UCR3 a D12L genů rezistence jsou znázorněny. (C) Kapkový test. Buňky uvedených kmenů byly kultivovány v YPD médiu při 28 °C do  $OD_{600nm} = 1$  a 5  $\mu$ l kultury v uvedených ředěních bylo nakapáno na agarovou půdu SD– s G418 (100  $\mu$ g/ml). (D) Elektroforetogram detekce cDNA transkriptu genu *G418<sup>R</sup>* pomocí PCR (35 cyklů) s primery kanR + KanVerF. (E) Elektroforetogram 5' RACE-PCR genu *G418<sup>R</sup>*. Specifický produkt u kontrolního kmene je označen vpravo. Šipka označuje specifický produkt vznikající z kryptického promotoru, umístěného někde v genu *K2ORF6*. M2: GeneRuler™ 100 bp Plus DNA Ladder.

**Závěr:** Výsledky ukazují, že vybraný promotor časného genu *D12L* viru *Vaccinia*, obsahující iniciátorový region, nefunguje ani jako slabý promotor v pGKL plasmidech. Ačkoliv v připraveném kmeni *K. lactis* IFO1267\_pRK2-27 dochází k slabé expresi genu *G418<sup>R</sup>*, jež se dá detekovat fenotypově a na úrovni RNA, transkripty *G418<sup>R</sup>* nevznikají z přilehlého virového promotoru *D12L*, ale z distálního kryptického plasmidového promotoru, lokalizovaného přibližně ve třetině genu *K2ORF6*.

### 5.1.15 Definovaná tvorba 3' konců plasmidových mRNA není závislá na promotoru a kódující sekvenci genu *in vivo*

Dosavadní publikované výsledky naznačují, že u pGKL plasmidů pravděpodobně existuje definovaný mechanismus terminace transkripce využívající jednosměrné terminátory (STARK *et al.* 1990; SCHAFFRATH *et al.* 1995b; JESKE *et al.* 2006). Mapování 3' konců mRNA stejně jako detekce možných posttranskripčních modifikací 3' konců mRNA pGKL plasmidů však zatím nebylo úspěšné, ačkoliv transkripty plasmidu pGKL1 se vážou na oligo(dT)-celulózovou kolonu, což by teoreticky mohlo nasvědčovat 3' polyadenylaci mRNA pGKL plasmidů (SOR a FUKUHARA 1985; STARK *et al.* 1990). V naší laboratoři se díky experimentům Václava Vopálenského využívajícím 3' RACE-PCR podařilo detekovat 3' konce mRNA všech otevřených čtecích rámců pGKL plasmidů. Bylo prokázáno, že 3' konce mRNA nejsou 3' polyadenylované *in vivo*. Sekvence jednotlivých cDNA klonů těchto experimentů jsem měl k dispozici pro bioinformatickou analýzu (~22 sekvencí jednotlivých cDNA klonů pro každý ORF; 344 sekvencí celkem).

Nejprve jsem se snažil manuálně a pomocí programu MEME 4.11<sup>4</sup> identifikovat případný sekvenční element v blízkosti detekovaných 3' konců mRNA, který by se vyskytoval u všech otevřených čtecích rámců pGKL plasmidů, ale žádný jsem nenašel. Poté jsem se snažil manuálně detekovat transkripční terminační signál 5'-TTTTTNT-3', vyskytující se u časných genů VACV ve vzdálenosti 100–30 nukleotidů proti směru transkripce od detekovaných 3' konců mRNA. Tento sekvenční motiv se mi podařilo identifikovat v blízkosti 3' konců mRNA pouze u genů *K1ORF1*, *K2ORF1* a *K2ORF7*. Tento motiv se ovšem v pGKL plasmidech vyskytuje často (celkově 116x) a jeho umístění se zdá na první pohled náhodné, takže tento motiv s největší pravděpodobností není terminačním signálem transkripce pGKL plasmidů. Rozhodl jsem se otestovat, zda opravdu dochází k definované tvorbě 3' konců plasmidových mRNA s využitím již připravených kvasinkových kmenů *K. lactis* s rekombinantními pGKL plasmidy, kde jsem detekoval 3' konce mRNA vnesených genů rezistencí ke G418 a hygromycinu B.

<sup>4</sup> <http://meme-suite.org/tools/meme>

Pro 3' RACE-PCR analýzu jsem vybral kmeny *K. lactis* IFO1267\_pRKL2-3 (Flag-K2ORF6p) a IFO1267\_pRKL2-12 (HA-K2ORF6p), kde se geny *G418<sup>R</sup>* a *HygB<sup>R</sup>* nacházejí před genem *K2ORF6* ve stejné transkripční orientaci, jako gen *K2ORF5* před genem *K2ORF6* v kmeni *K. lactis* IFO1267 (Obrázek 48A). Na provedení 3' RACE-PCR experimentů, izolaci specifických produktů PCR a jejich klonování do vektorů, analýze vektorů pomocí restričního štěpení, izolaci vektorů a jejich sekvenaci se podílela Jitka Vojáčková.

Z uvedených kvasinkových kmenů byla izolována celková RNA (metoda 4.2.3.1.1). Zbývající DNA ve vzorcích byla odstraněna deoxyribonukleázou I (metoda 4.2.3.1.4) a kvalita RNA byla ověřena pomocí agarosové elektroforézy (metoda 4.2.3.3.1) a UV spektrofotometrie (4.2.3.2). Purifikovaná celková RNA byla použita k syntéze homopolymerního 3' konce (metoda 4.2.3.4.2) a posléze k reverzní transkripci využívající primer oligo(dG)anch2 (metody 4.2.3.6.1 a 4.2.3.6.3). Vzniklá cDNA byla purifikována (metoda 4.2.2.1.7) a použita pro 3' RACE-PCR s primery uvedenými v Tabulce 28. Produkty 3' RACE-PCR reakcí byly rozděleny agarosovou elektroforézou (metoda 4.2.2.3.1), specifické produkty byly izolovány (metoda 4.2.2.1.5) a klonovány do vektoru pCR™4-TOPO® (metoda 4.2.2.4.5). Vektory byly transformovány do buněk *E. coli* XL-1 Blue (metoda 4.2.2.5.1). Ze získaných monoklonií byly vektory následně izolovány (metoda 4.2.2.1.1) a charakterizovány pomocí restričního štěpení enzymy, uvedenými v Tabulce 28. Z kolonií, jejichž vektory vykazovaly přítomnost insertu odpovídajícímu specifickému produktu 3' RACE-PCR reakce, byly vektory izolovány (metoda 4.2.2.1.2), jejich kvalita byla ověřena UV spektrofotometrií (4.2.2.2), a následně byly podrobeny sekvenaci DNA (metoda 4.2.2.8.4) v oblasti insertu primerem T7. Elektroforetogramy detekce produktů 3' RACE-PCR a sekvence jednotlivých klonů jsou na Obrázku 48.

**Tabulka 28: 3' RACE-PCR analýza transkriptů pGKL plasmidů**

Gen	Promotor	Kmen	Primery pro 3' RACE-PCR (35 cyklů)	Produkty	Klonováno do vektoru	Restriční analýza vektoru	Sekvenace vektoru
<i>K2ORF5</i>	K2UCR5	IFO1267	in_ORF5_forw anch2	643 bp 615 bp	pCR™4-TOPO®	<i>EcoRI</i> <i>VspI</i>	T7
<i>G418<sup>R</sup></i>	K1UCR2	IFO1267_pRKL2-3	KanVerF anch2	387 bp 415 bp	pCR™4-TOPO®	<i>EcoRI</i> <i>NcoI</i>	T7
<i>HygB<sup>R</sup></i>	K1UCR3	IFO1267_pRKL2-12	Hygro_tail anch2	418 bp 446 bp	pCR™4-TOPO®	<i>EcoRI</i> <i>NcoI</i>	T7

Dřívější výsledky V. Vopálenského na kmeni *K. lactis* IFO1267 ukázaly, že pro transkripty genu *K2ORF5* je pomocí 3' RACE-PCR možno detekovat dvě populace s rozdílnou délkou 3' konců mRNA lišící se přibližně o 28 nukleotidů (data nezobrazena). Výsledky 3' RACE-PCR genů *G418<sup>R</sup>* a *HygB<sup>R</sup>*, které jsou v rekombinantních kmenech lokalizovány vzhledem ke genu *K2ORF6* ve stejné orientaci jako *K2ORF5* ke genu *K2ORF6* v kmeni IFO1267, ukázaly, že 3' konce mRNA těchto genů se v závislosti na délce také dají rozdělit na dvě délkové



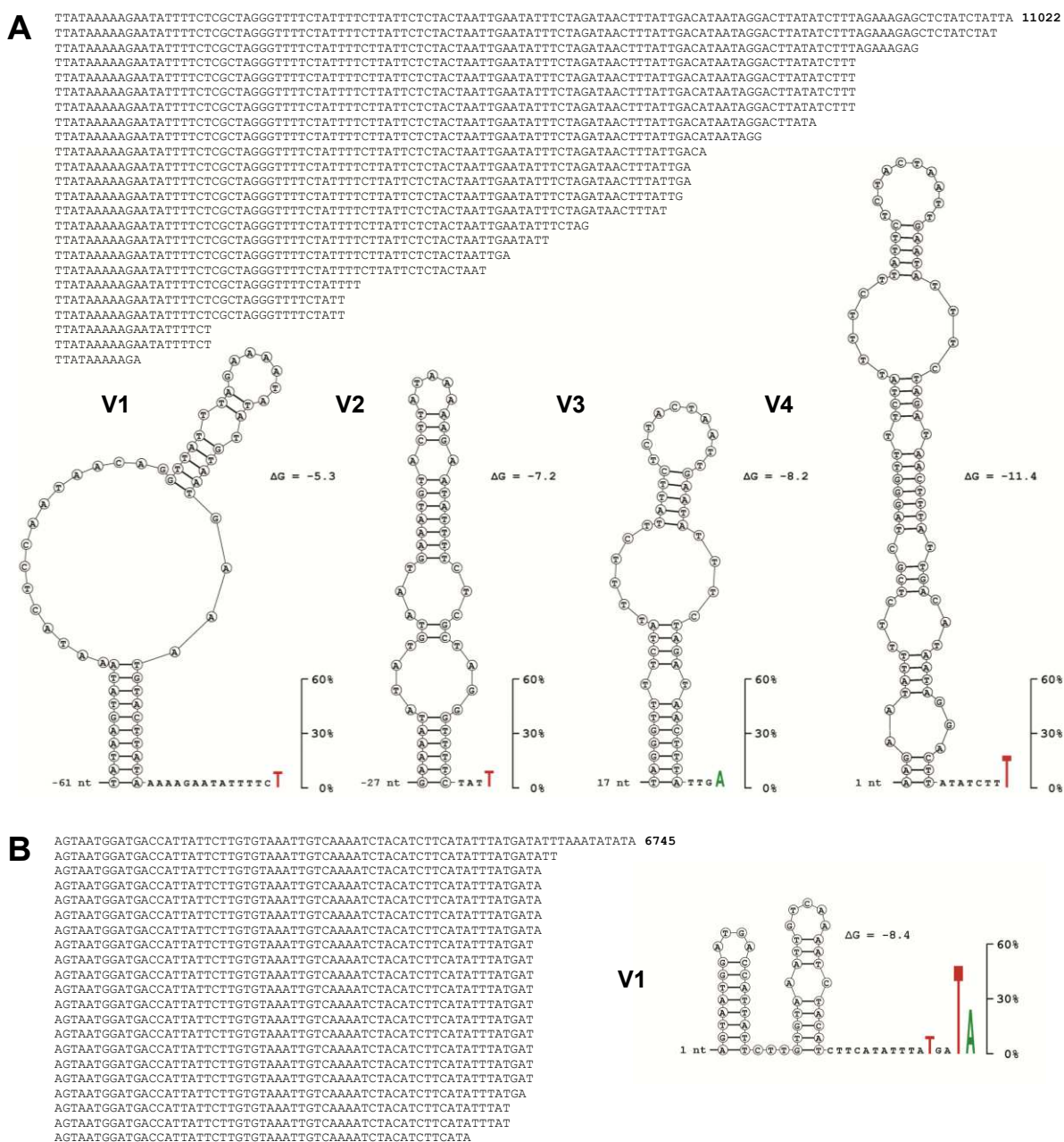
### 5.1.16 3' konce plasmidových mRNA se nachází poblíž předpokládaných RNA vlásenek

Protože výsledky předchozí kapitoly jasně poukazují na definovaný mechanismus tvorby 3' konců mRNA pGKL plasmidů, rozhodl jsem se opětovně bioinformaticky analyzovat sekvence jednotlivých cDNA klonů 3' RACE-PCR experimentů, pokrývajících celý genom pGKL plasmidů, získané V. Vopálenským. Jako klíčková událost tohoto snažení byla konzultace zmíněných sekvencí s Liborem Krásným, který mě nasměroval na *in silico* analýzu sekundárních struktur RNA poblíž experimentálně určených 3' konců plasmidových mRNA. Pomocí serveru RNAstructure (REUTER a MATHEWS 2010) se mi podařilo identifikovat předpokládané struktury RNA vlásenek poblíž 3' konců mRNA u všech genů pGKL plasmidů. Předpokládané RNA vlásenky začínají typicky relativně blízko za příslušným stop kodónem daného otevřeného čtecího rámce (medián vzdálenosti 26 nt) a jejich konce jsou v blízkosti detekovaných 3' terminálních nukleotidů mRNA (medián vzdálenosti 15 nt). Gibbsova volná energie předpokládaných RNA vlásenek je však relativně vysoká (medián  $-7,5$  kcal/mol) viz Tabulka 29.

**Tabulka 29:** Předpokládané RNA vlásenky poblíž 3' konců mRNA pGKL plasmidů

Gen	Vlásenka	Vzdálenost k stop kodónu	$\Delta G$	Nukleotidová sekvence 5' → 3'
<i>K1ORF1</i>	V1	-12	-4,6	GTGAAAGCTACTTAATATATGAAAGTTTTATAATAATTATAAAATGAATATATTTTA
<i>K1ORF2</i>	V1	1	-8,4	AGTAATGGATGACCATTATTCTTGTTGTAAATTGTCAAATCTACAT
<i>K1ORF3</i>	V1	99	-7,8	TGTTTTCTCTAAGAAGCAAATACTGGTAAGAATAAACCTGCTAAAGCGAATGCTAGGT
	V2	145	-12,7	GCGAATGCTAGGTCTGATGGATTAACTCTCCGTTTATAGCGGCATCTG
<i>K1ORF4</i>	V1	127	-3,6	TGGAGACGTACTAATCTTCCA
<i>K2ORF1</i>	V1	230	-6,4	GCTTATTATAGCATTCTTAGTATCGTCTATTGTACTACTCTAATTGTCTAATTCGC
<i>K2ORF2</i>	V1	110	-8,7	TGGATTAAATTTAGGTTCTATATCAGGAATTTCTTCTATACCTGGTTCTTCTTAACAATAATTATCATAA TCTA
<i>K2ORF3</i>	V1	-22	-8,0	TTATCATTTCTTTTCTAAAAAATGAATGATGAATTAGCATTCTTAAATCTCAAATAGATGACTA
	V2	13	-3,3	ATTCTTAAATCTCAAATAGATGACTATTCCGATGATATAGAAT
	V3	58	-3,0	TCTAAGAACAGAAGTATTTAGA
	V4	73	-6,4	ATTTAGAATAAATAGATATTCTAAAT
<i>K2ORF4</i>	V1	9	-8,7	CACCAAATTAATTAATTAGGTAAAACTAATAAATTTAGTGGTG
	V2	45	-9,1	TAGTGGTGAAGAGGAAAAATCTATTCTTCTTTATACTA
<i>K2ORF5</i>	V1	118	-4,8	AAGTTACTGTGCTATATACAATGACAGAGAAGCTT
	V2	142	-8,6	CAGAGAAGTTCTATAATGATTGGTAGTTCACTG
<i>K2ORF6</i>	V1	-61	-5,3	TATAAGTATAAATACTCCAATAACAGGTTATTTAGAAAAATATATGTAATGAAATGTACTT ATA
	V2	-27	-7,2	GAAAAATATATGTAATGAAATGTACTTATAAAAAAGAATATTTCTCGCTAGGGTTTTTC
	V3	17	-8,2	TAGGGTTTTCTATTTCTTATTCTCTACTAATTGAATATTTCTAGATAACTTTA
	V4	1	-11,4	AAGAATATTTTCTCGCTAGGGTTTTCTATTTTCTTATTCTCTACTAATTGAATATTTCTAGATAACTTTATTGA CATAATAGGACTT
<i>K2ORF7</i>	V1	-4	-8,0	TTTTATAAGTACATTTTATTACATATATTTTCTAAATAACCTGTTATTGGAGTATTTATACTTATAGA
<i>K2ORF8</i>	V1	-12	-4,3	GATGATATAAAGTAAAAATGAATGAAAAATTTATTTCTAATTTAACTATTTCTCAGAGATAACCCAGAGA ACTTTAAGT
<i>K2ORF9</i>	V1	35	-2,0	CGATCAAAAAAGATTATTTAAAAAGAAATAATAAAGAAAAATTG
<i>K2ORF10</i>	V1	-19	-6,9	AGTAGACATTAACAATTTAAATTTATATGTAACATTAATATACATTTAACACTGTATGCTACT
<i>K2ORF11</i>	V1	108	-15,8	TTATGATGTTGGTTTAGGTAAAAACATATGCTGCTGCATGCTTGCACATATGTATCTAGATTCTGGTTTTAA



[illegible]

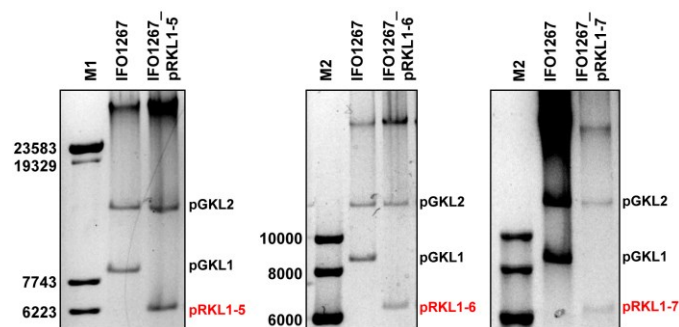
166



**Závěr:** Výsledky ukazují, že 3' konce transkriptů pGKL plasmidů se nachází poblíž předpokládaných struktur slabých RNA vlásenek. Tyto struktury by se tak možná mohly podílet na definované tvorbě 3' konců mRNA pGKL plasmidů. Ve většině případů se předpokládané RNA vlásenky nacházejí v 3' nepřekládané oblasti transkriptu až za stop kodónem příslušného otevřeného čtecího rámce, což je v souladu s mým předchozím zjištěním, že definovaná tvorba 3' konců plasmidových transkriptů není závislá na promotoru ani na sekvenci a délce kódující oblasti genu příslušného transkriptu.

### 5.1.17 RNA vlásenky se podílí na tvorbě 3' konců plasmidových mRNA *in vivo*

Rozhodl jsem se otestovat, zda předpokládané struktury RNA vlásenek ovlivňují tvorbu 3' konců mRNA pGKL plasmidů. Protože pGKL plasmidy neobsahují téměř žádné intergenové oblasti, předpokládané RNA vlásenky jsou lokalizovány v kódujících sekvencích sousedních otevřených čtecích rámců nebo v terminálních invertovaných repetících. To znamená, že sekvence těchto RNA vlásenek nemohou být podrobeny mutagenезi bez možnosti změny plasmidových funkcí. Proto byly připraveny tři kmeny *K. lactis* s modifikovanými pGKL plasmidy postupem detailněji popsáním v kapitole 5.1.1.1. Připravené kmeny obsahovaly modifikovaný plasmid pGKL1 kódující gen *G418<sup>R</sup>* (místo genu *K1ORF2*) pod kontrolou promotoru K1UCR2, po kterém následovala 3' nepřekládaná oblast genu *K2ORF5* divokého typu (IFO1267\_pRKL1-5), 3' nepřekládaná oblast genu *K2ORF5* s mutacemi destabilizujícími RNA vlásenku V2 (IFO1267\_pRKL1-6) a 3' nepřekládaná oblast genu *K2ORF5* s mutacemi obnovujícími párování bází v RNA vlásece V2 (IFO1267\_pRKL1-7). Elektroforetogramy detekce rekombinantních pGKL plasmidů v připravených kmenech jsou na Obrázku 50. Schéma kódující kapacity rekombinantních pGKL plasmidů je na Obrázku 51A. Nejdůležitější informace k přípravě a charakterizaci nových kmenů souvisejících s touto kapitolou jsou shrnuty v Tabulce 30.



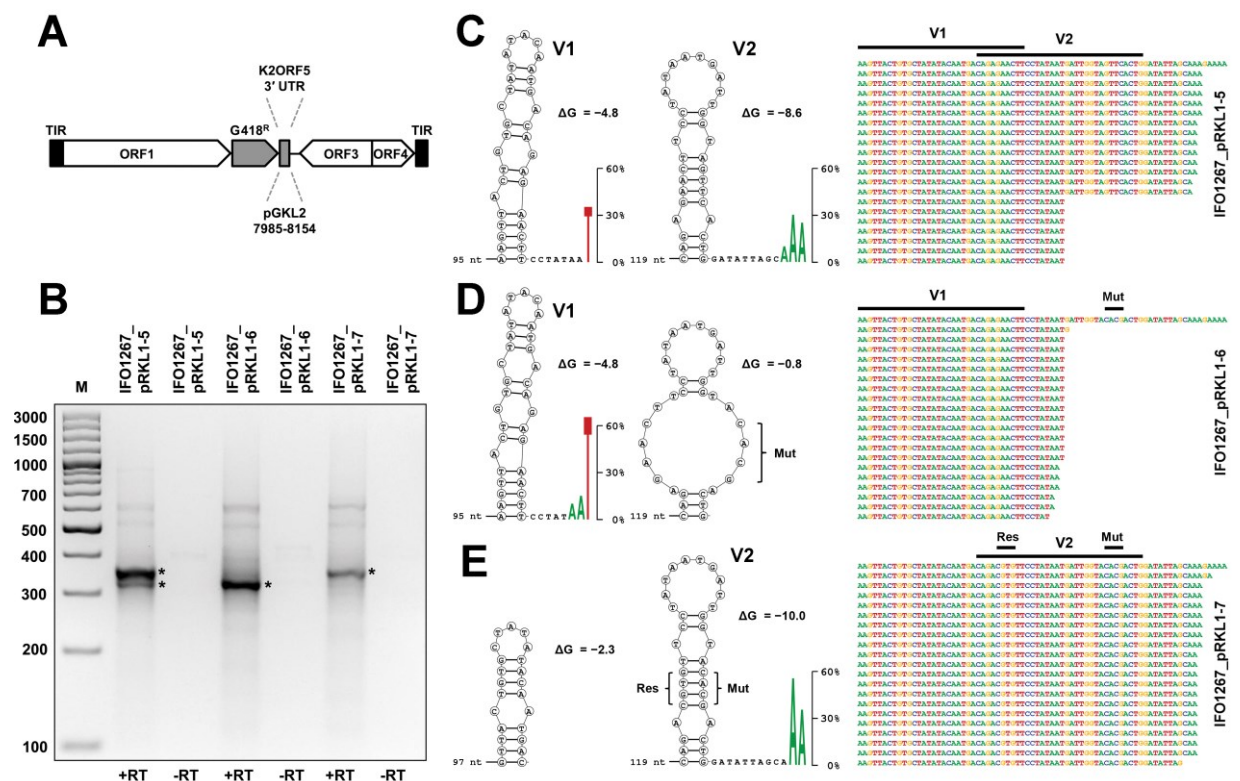
**Obrázek 50:** Elektroforetogramy DNA připravených kmenů kódujících gen *G418<sup>R</sup>* pod kontrolou promotoru K1UCR2, následovaný různými variantami 3' nepřekládané oblasti (UTR) genu *K2ORF5*. Rekombinantní pGKL plasmidy jsou označeny červeně. Velikost rekombinantních pGKL plasmidů je uvedena v Tabulce 6. Zobrazeny jsou vybrané monoklonie po pěti pasážích nezávislých klonů získaných po transformaci. M1: Lambda DNA/Eco130I (*Sfi*I) Marker, 16; M2: GeneRuler™ 1 kb DNA Ladder.

**Tabulka 30:** Primery a templáty pro přípravu a charakterizaci kmenů této kapitoly

Plasmid	Primery na přípravu kazety pomocí PCR	Templát pro přípravu kazety	Primery na přípravu kazety pomocí fúzní PCR	Primery pro ověření integrace kazety a přípravu templátu na sekvenování / Sekvenace	Primery pro ověření integrace kazety 2
pRKL1-5	pGKL-kanF kanR2	pRKL1-1	pGKL-kanF K1_K2O5-UTR_2R	KanVerF ORF3-pGKL1_tail	in_Kan_rev1 ORF1-K1_tail_2
	K1_K2O5-UTR_2F K1_K2O5-UTR_2R	pGKL2		KanVerF	
pRKL1-6	pGKL-kanF kanR2	pRKL1-1	pGKL-kanF K1_K2O5-UTRm_2R	KanVerF ORF3-pGKL1_tail	in_Kan_rev1 ORF1-K1_tail_2
	K1_K2O5-UTR_2F K1_K2O5-UTRm_2R	pGKL2		KanVerF	
pRKL1-7	pGKL-kanF kanR2	pRKL1-1	pGKL-kanF K1_K2O5-UTRres_2R	KanVerF ORF3-pGKL1_tail	in_Kan_rev1 ORF1-K1_tail_2
	K1_K2O5-UTR_2F K1_K2O5-UTRres_2R	pGKL2		KanVerF	

Z uvedených kvasinkových kmenů byla izolována celková RNA (metody 4.2.3.1.1 a 4.2.3.1.2). Zbývající DNA ve vzorcích byla odstraněna deoxyribonukleázou I (metoda 4.2.3.1.4) a kvalita RNA byla ověřena pomocí agarosové elektroforézy (metoda 4.2.3.3.1) a UV spektrofotometrie (4.2.3.2). Purifikovaná celková RNA byla použita k syntéze homopolymerního 3' konce (metoda 4.2.3.4.2) a posléze k reverzní transkripci využívající primer oligo(dG)anch2 (metody 4.2.3.6.1 a 4.2.3.6.3). Vzniklá cDNA byla purifikována (metoda 4.2.2.1.7) a použita pro 3' RACE-PCR s primery KanVerF a anch2. Produkty 3' RACE-PCR reakcí byly rozděleny agarosovou elektroforézou (metoda 4.2.2.3.1), specifické produkty byly izolovány (metoda 4.2.2.1.5) a klonovány do vektoru pCR<sup>™</sup>4-TOPO<sup>®</sup> (metoda 4.2.2.4.5). Vektory byly transformovány do buněk *E. coli* XL-1 Blue (metoda 4.2.2.5.1). Ze získaných monoklonií byly vektory následně izolovány (metoda 4.2.2.1.1) a charakterizovány pomocí restriční štěpení enzymy *EcoRI* a *NcoI*. Z kolonií, jejichž vektory vykazovaly přítomnost insertu odpovídajícímu specifickému produktu 3' RACE-PCR reakce, byly vektory izolovány (metoda 4.2.2.1.2), jejich kvalita byla ověřena UV spektrofotometrií (4.2.2.2), a následně byly podrobeny sekvenaci DNA (metoda 4.2.2.8.5) v oblasti insertu primerem T7. Elektroforetogramy detekce produktů 3' RACE-PCR jsou na Obrázku 51B. Sekvence jednotlivých cDNA klonů a předpokládané RNA vlásenky v daných kmenech jsou zobrazeny na Obrázku 51C, D a E.

Distální část 3' UTR *K2ORF5* obsahuje dvě předpokládané částečně se překrývající RNA vlásenky nazvané V1 a V2 (Obrázek 51C). Výsledky 3' RACE-PCR experimentů z kmene *K. lactis* IFO1267\_pRKL1-5 ukázaly, že transkripty *G418<sup>R</sup>* lze v závislosti na délce 3' UTR rozdělit na dvě populace, přičemž terminální nukleotidy se nacházejí poblíž RNA vlásenky V1 a V2 (Obrázek 51B a C).



**Obrázek 51:** RNA vlásenky se podílí na tvorbě 3' konců plasmidových mRNA *in vivo*. **(A)** Schematické znázornění rekombinantního plasmidu pGKL1, kde je gen *G418<sup>R</sup>* následován kódující sekvencí 3' UTR genu *K2ORF5* divokého typu (pRKL1-5) či modifikovanou (pRKL1-6, pRKL1-7). TIR: terminální invertované repetice. **(B)** 3' RACE-PCR analýza mRNA genu *G418<sup>R</sup>*, exprimovaného z modifikovaných plasmidů pGKL1. M: GeneRuler™ 100 bp Plus DNA Ladder. Specifické produkty, které byly klonovány do vektoru pCR™4-TOPO® a použity pro sekvenování jsou označeny hvězdičkami. Reverzní transkripce byla prováděna za přítomnosti (+RT) a nepřítomnosti (–RT) reverzní transkriptázy. **(C)** 3' RACE-PCR analýza genu *G418<sup>R</sup>* z kmene *K. lactis* IFO1267\_pRKL1-5. V tomto a následujících panelech odpovídá pravá horní sekvence DNA templátu (plasmid); sekvence umístěné níže představují jednotlivé sekvenované cDNA klony (zobrazen je pouze konec 3' nepřekládané oblasti). Umístění předpokládaných RNA vlásenek je vyznačeno nad sekvencemi. RNA vlásenky jsou zobrazeny jako písmena cDNA nukleotidů v kruzích a pro každou strukturu jsou zobrazeny hodnoty Gibbsovy volné energie (ΔG) v kcal/mol. Vzdálenost začátku RNA vlásenek od stop kodónu genu příslušného transkriptu je uvedena jako počet nukleotidů (nt). Terminální nukleotidy experimentálně stanovených 3' konců cDNA jsou znázorněny jako barevná písmena zvětšená proporcionálně k jejich výskytu (v procentech) v sekvenovaných klonech, když byl detekován stejný terminální nukleotid u sekvencí alespoň dvou nezávislých klonů. **(D)** 3' RACE-PCR analýza genu *G418<sup>R</sup>* z kmene *K. lactis* IFO1267\_pRKL1-6. Mut: mutace ve stonku RNA vlásenky V2. **(E)** 3' RACE-PCR analýza genu *G418<sup>R</sup>* z kmene *K. lactis* IFO1267\_pRKL1-7. Mut: mutace ve stonku RNA vlásenky V2. Res: mutace obnovující párování bází ve stonku RNA vlásenky V2.

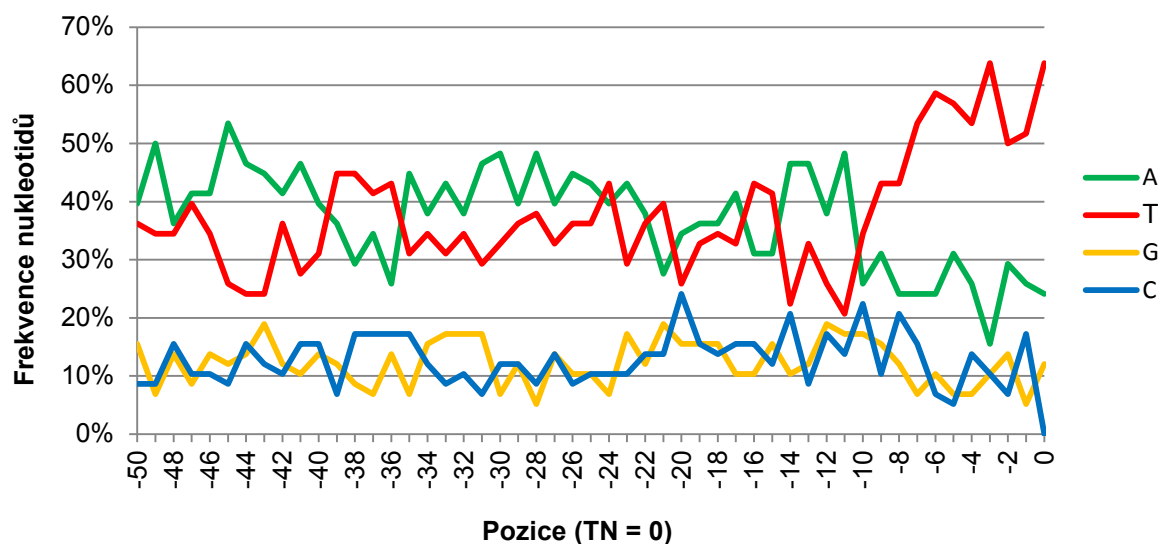
Kmen *K. lactis* IFO1267\_pRKL1-6 obsahoval mutace 4 nukleotidů destabilizující párování bází uprostřed stonku předpokládané RNA vlásenky V2 (Obrázek 51D). Výsledky 3' RACE-PCR experimentů z tohoto kmene ukázaly jednu délkovou populaci transkriptů *G418<sup>R</sup>*, přičemž terminální nukleotidy se nacházejí poblíž RNA vlásenky V1 (Obrázek 51B a D). Následně byl připraven kmen *K. lactis* IFO1267\_pRKL1-7, kde byla změněná sekvence,

nikoliv však párování bází 4 nukleotidů uprostřed stonku předpokládané RNA vlásenky V2 (Obrázek 51E). Kompenzační mutace předpokládané RNA vlásenky V2 zároveň výrazně změnily délku a Gibbsovu volnou energii překryvné RNA vlásenky V1 (Obrázek 51E). Výsledky 3' RACE-PCR experimentů z tohoto kmene ukázaly jednu populaci transkriptů *G418<sup>R</sup>*, přičemž terminální nukleotidy se nacházejí poblíž obnovené RNA vlásenky V2 (Obrázek 51B a E). Destabilizace RNA vlásenky V1 v kmeni IFO1267\_pRKL1-7 je pravděpodobně důvodem, proč nebyly detekovány transkripty s terminálními nukleotidy poblíž pozměněné RNA vlásenky V1 (Obrázek 51B a E).

**Závěr:** Výsledky ukazují, že předpokládané RNA vlásenky jsou esenciální pro definovanou tvorbu 3' konců transkriptů pGKL plasmidů *in vivo*. Předpokládané RNA vlásenky by mohly fungovat jako terminátory v mechanismu vnitřní terminace transkripce, jenž kromě RNA polymerázy nevyžaduje přídatné proteinové faktory.

#### 5.1.18 Transkripty pGKL plasmidů obsahují předpokládaný 3' oligo(U) konec

Struktury RNA vlásenek vnitřních terminátorů transkripce u bakterií jsou obvykle následovány typickým 7–8 nukleotidů dlouhým 3' oligo(U) koncem, který podporuje zastavení elongující RNA polymerázy na slabé DNA:RNA hybridní molekule obsahující dA:rU nukleotidové páry (MARTIN a TINOCO 1980; D'AUBENTON CARAFA *et al.* 1990; GUSAROV a NUDLER 1999). Je však známo, že 3' oligo(U) konec RNA není nezbytně nutný pro účinnou vnitřní terminaci transkripce u bakterií *in vitro* a *in vivo* (INGHAM *et al.* 1995; UNNIRAMAN *et al.* 2001). Sekvence jednotlivých cDNA klonů 3' RACE-PCR experimentů se na první pohled díky AT bohatosti pGKL plasmidů nezdají bohaté na thymidinové nukleotidy na svých 3' koncích. Detailní bioinformatická analýza nukleotidového složení 3' konců mRNA pGKL plasmidů však odhalila nabohacení T nukleotidů v posledních 8 nukleotidech 3' konců cDNA pGKL plasmidů, což odpovídá předpokládanému 3' oligo(U) konci mRNA (Obrázek 52).



**Obrázek 52:** Transkripty pGKL plasmidů obsahují předpokládaný 3' oligo(U) konec. Tento obrázek představuje výsledky sekvenční analýzy všech klonů cDNA, jejichž terminální nukleotid byl detekován alespoň u dvou nezávislých klonů, jejichž sekvence byly získány 3' RACE-PCR analýzou mRNA odpovídající všem genům kódovaným pGKL plasmidy. Tyto cDNA sekvence (58 unikátních sekvencí) byly ručně zarovnány podle svých terminálních nukleotidů (TN), frekvence jednotlivých nukleotidů byla vypočtena pro každou pozici 50 nukleotidů proti směru transkripce od terminálního nukleotidu pomocí Seqool 3.1 a vynesena do grafu. Výskyt thymidinových zbytků v 8 terminálních nukleotidech 3' konců cDNA (včetně TN) je  $\geq 50\%$ .

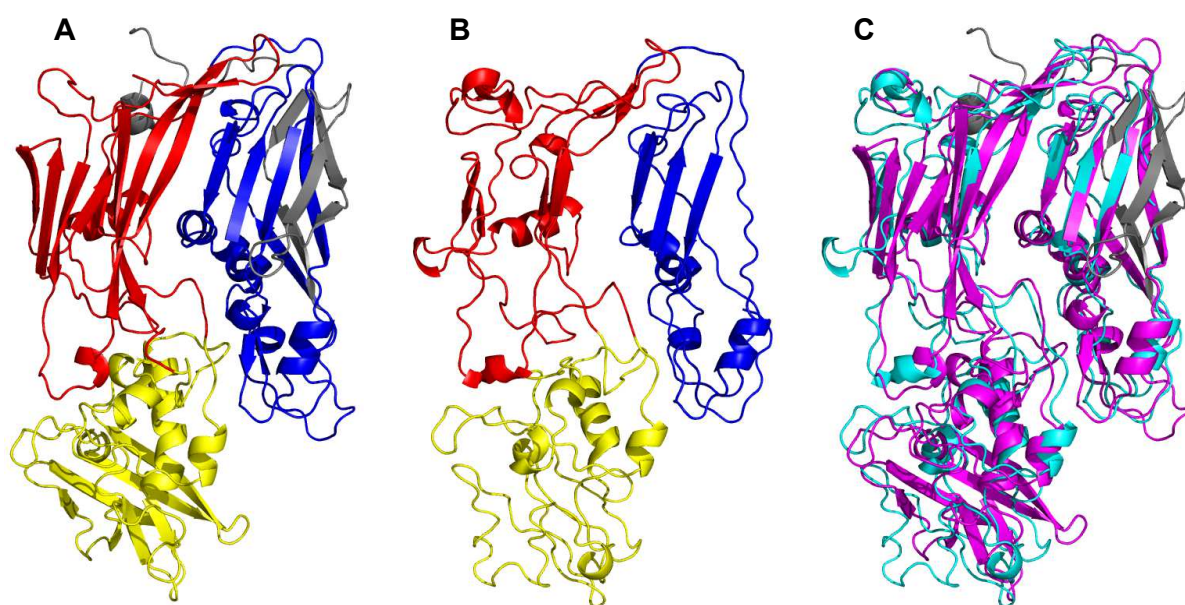
**Závěr:** Výsledky ukazují, že transkripty pGKL plasmidů obsahují předpokládaný 3' oligo(U) konec, který by se mohl spolu s RNA vlásenkami podílet na definované tvorbě 3' konců transkriptů pGKL plasmidů mechanismem vnitřní terminace transkripce.

### 5.1.19 Gen *K2ORF9* pGKL plasmidů kóduje předpokládaný *scaffold* protein

Z 15 otevřených čtecích rámců pGKL plasmidů dosud nebyla přisouzena role proteinovým produktům genů *K2ORF1*, *K2ORF8*, *K2ORF9* a *K2ORF11*. Dalším cílem této práce proto bylo pokusit se určit identitu plasmidových otevřených čtecích rámců s dosud neznámou funkcí, především za účelem zjištění, zda se některý z těchto proteinů také nepodílí na procesu transkripce pGKL plasmidů. Náš zájem o dosud necharakterizované ORF plyne především z toho, že se navzdory několika nezávislým pokusům nepodařilo prokázat methyltransferázovou aktivitu mRNA *capping* enzymu K2ORF3p *in vitro* (TIGGEMANN *et al.* 2001; VOPÁLENSKÝ 2007), ačkoliv naše výsledky naznačují, že transkripty pGKL plasmidů nesou N<sup>7</sup>-methylguanosinovou čepičkovou strukturu. Některý z otevřených čtecích rámců s neznámou funkcí by tak mohl kódovat druhou podjednotku mRNA *capping* enzymu stimulující methyltransferázovou aktivitu K2ORF3p, jak je tomu u mRNA *capping* enzymu viru *Vaccinia*. Použil jsem proto sekvence



aminokyselinových zbytků produktů genů *ORF1*, *ORF8*, *ORF9* a *ORF11* plasmidu pGKL2 a sekvence produktů homologních genů u kvasinkových cytoplasmatických plasmidů pSKL, pPac1-1 a pPE1B pro hledání sekvenčně podobných proteinů v dostupných databázích různými bioinformatickými přístupy. Jediný potenciálně biologicky relevantní výsledek tohoto snažení byl získán modelováním terciárních struktur proteinů neznámých funkcí serverem Robetta, přičemž pomocí této strategie se mi již dříve podařilo detekovat přítomnost dosud nenalezených konzervovaných oblastí  $\beta$ a1,  $\beta$ a6 a  $\beta$ a13  $\beta$  podjednotky RNA polymerázy v proteinu K2ORF6p (SÝKORA 2013). Při použití sekvencí produktů genu *ORF9* plasmidu pGKL2 *K. lactis* a *ORF9* plasmidu pSKL *Saccharomyces kluyveri* pro tvorbu homologních modelů terciárních struktur pomocí serveru Robetta byly oba tyto proteiny shodně modelovány podle proteinu D13 viru *Vaccinia*. Protein D13 VACV je takzvaný *scaffold* protein, který trimerizuje a tvoří vnější mřížovaný povrch nezralých sférických částic viru (SZAJNER *et al.* 2005). Protein D13, fylogeneticky patřící do rodiny kapsidových proteinů ikosahedrálních DNA virů, je schopný remodelace lipidových membrán (BAHAR *et al.* 2011; HYUN *et al.* 2011). Protein D13, jehož terciární struktura je na Obrázku 53A, obsahuje 551 aminokyselinových zbytků. Proteiny K2ORF9p a pSKL-ORF9p obsahují 453 a 458 aminokyselinových zbytků, přičemž modely těchto struktur jsou celkově strukturám proteinu D13 velmi podobné a kromě částečné redukce postrádají oblast odpovídající 75 aminokyselinovým zbytkům N-konce proteinu D13. Model proteinu K2ORF9p a jeho proložení se strukturou proteinu D13 je na Obrázku 53B a C.



**Obrázek 53:** Protein K2ORF9p je předpokládaný homolog *scaffold* proteinu viru *Vaccinia*. (A) Struktura proteinu D13 viru *Vaccinia* (PDB ID: 2YGC.B). N-koncová část chybějící u modelů K2ORF9p a pSKL-ORF9p zobrazena šedě. Dvě  $\beta$ -barelové roládové domény zobrazeny modře a červeně. Hlavová doména zobrazena žlutě. Domény jsou označeny podle (HYUN *et al.* 2011). (B) Model proteinu K2ORF9p. (C) Vzájemné proložení struktury proteinu D13 (N-konec šedě, zbytek fialově) a modelu proteinu K2ORF9p (tyrkysově). Strukturální porovnání má hodnotu RMSD 3,2 Å přes 406 aminokyselinových zbytků. Struktury byly porovnány pomocí DaliLite v. 3 a vizualizovány v PyMOL 1.3.



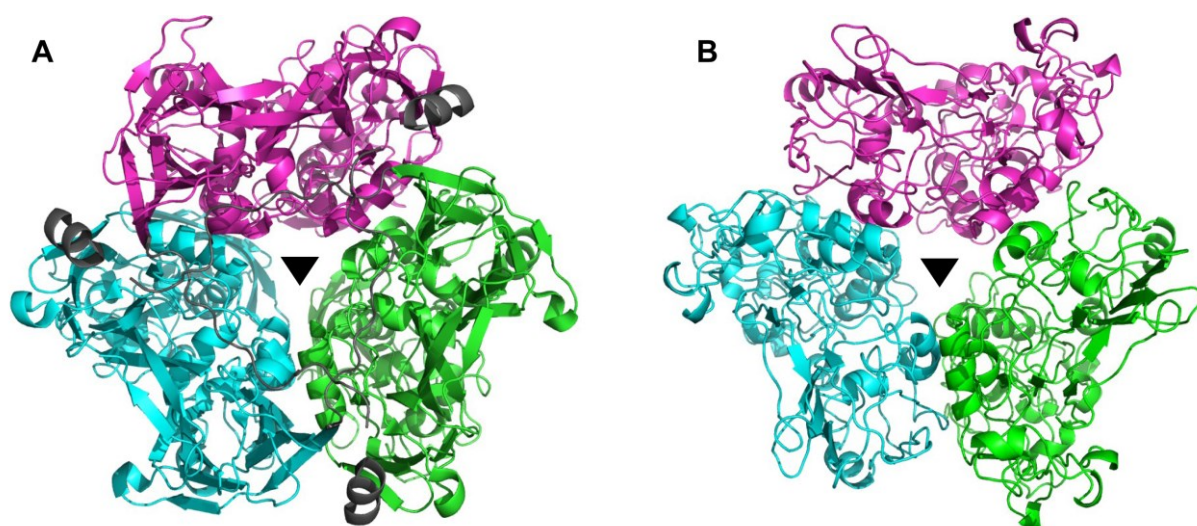
**Závěr:** Geny *ORF9* kvasinkových cytoplasmatických lineárních plasmidů kódují předpokládaný *scaffold* protein, homologní k proteinu D13 viru *Vaccinia*. Tento protein by se teoreticky mohl podílet na formaci vnitrobuněčných membránových struktur kolem genomu pGKL plasmidů, podobných nezralým částicím viru *Vaccinia*.

### 5.1.20 Gen *D13L* viru *Vaccinia* nedokáže zastoupit funkci genu *K2ORF9* pGKL plasmidů *in vivo*

Celková sekvenční identita aminokyselinových zbytků strukturního porovnání proteinu D13 a modelu K2ORF9p je pouhých 10 %, což je poměrně nízká hodnota. Proto se mi pravděpodobně nepodařilo detekovat sekvenční podobnost těchto dvou proteinů pomocí BLASTP® hledání či hledání využívající skrytých Markovových modelů. Nízká sekvenční identita obou proteinů znamená, že pouze na základě sekvence může být podobnost těchto proteinů čistě náhodná. Z 11 otevřených čtecích rámců pGKL plasmidů, jejichž produkty mají předpovězenou funkci, se v naprosté většině jedná o geny enzymů a proteinů interagujících s nukleovými kyselinami. V takovém případě jsou požadavky na přítomnost definovaných aminokyselinových zbytků v aktivních místech a vazebných motivech těchto proteinů striktní, což způsobuje větší sekvenční identitu těchto proteinů s proteiny v databázích, a tedy i jednodušší identifikaci homologních proteinů. Protein D13 je však strukturní protein, u kterého jsou evolucí kladeny nároky spíše na zachování celkové struktury než na sekvence aminokyselinových zbytků, což dokládá i značná sekvenční odlišnost proteinu D13 od homologních kapsidových proteinů ze stejné proteinové rodiny (IYER *et al.* 2001; BAHAR *et al.* 2011).

Protein D13 VACV tvoří trimery (Obrázek 54A), které se váží na lipidové membrány, zakřivují je a odlamují, čímž se tvoří pŕlměsíité membránové struktury skládající se kolem genomu VACV (HEUSER 2005; SZAJNER *et al.* 2005; CHLANDA *et al.* 2009; HYUN *et al.* 2011). Z modelu proteinu K2ORF9p je patné, že strukturní prvky podílející se na vzájemných interakcích proteinů D13 v rámci trimery jsou u K2ORF9p zachovány (Obrázek 54B). Asociace proteinu D13 VACV s lipidovými membránami není spontánní a vyžaduje interakci proteinu D13 s N-koncem virového transmembránového proteinu A17 (BISHT *et al.* 2009; UNGER *et al.* 2013). Při maturaci nezralých částic VACV ve zralé virové částice dochází k disociaci proteinu D13 od membrán. Tato disociace je závislá na proteolytickém odštěpení N-konce proteinu A17 virovou proteázou I7 (ANSARAH-SOBRINHO a MOSS 2004; BISHT *et al.* 2009; UNGER *et al.* 2013). Předpokládá se, že by protein D13 mohl s virovým proteinem A17 interagovat skrze svou N-koncovou část obsahující amfipatický  $\alpha$ -helix (BAHAR *et al.* 2011), přičemž tuto N-koncovou

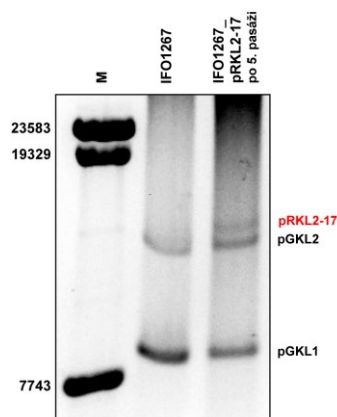
část proteiny K2ORF9p a pSKL-ORF9p postrádají. Rozhodl jsem se otestovat, zda by část genu *D13L* viru *Vaccinia*, kódující úsek proteinu D13, jehož část je pravděpodobně v proteinu K2ORF9p zachována, mohla zastoupit roli esenciálního genu *K2ORF9* pGKL plasmidů, čímž by se potvrdila domněnka o předpokládaném společném původu těchto genů. Proto byl připraven kmen *K. lactis* s modifikovaným pGKL plasmidem postupem detailněji popsáným v kapitole 5.1.1.1. Připravený kmen IFO1267\_pRKL2-17 obsahoval modifikovaný plasmid pGKL2 kódující místo genu *K2ORF9* část genu *D13L* viru *Vaccinia* kmene Western Reserve (aminokyselinové zbytky 76–551) pod kontrolou promotoru K2UCR9. Nejdůležitější informace k přípravě a charakterizaci tohoto kmene jsou shrnuty v Tabulce 31. Kmen *K. lactis* IFO1267\_pRKL2-17 se podařilo připravit, avšak rekombinantní plasmid pRKL2-17 nedokázal v buňkách nahradit plasmid pGKL2 při inkubaci v selektivních podmínkách favorizujících modifikovaný plasmid (Obrázek 55).



**Obrázek 54:** Protein K2ORF9p by mohl tvořit homotrimerní struktury. (A) Spodní pohled na strukturu trimeru proteinu D13 viru *Vaccinia* (PDB ID: 2YGC). Každý monomer je označen jinou barvou. N-koncový amfipatický helix každého monomeru, chybějící u modelů K2ORF9p a pSKL-ORF9p, je zobrazen šedě. Trojnásobná osa symetrie trimeru je označena černým trojúhelníkem. (B) Spodní pohled na model trimeru proteinu K2ORF9p. Každý monomer je označen jinou barvou, jako v případě Obrázku 54A. Struktury byly porovnány pomocí DaliLite v. 3 a vizualizovány pomocí PyMOL 1.3.

**Tabulka 31:** Primery a templáty pro přípravu a charakterizaci kmene této kapitoly

Plasmid	Primery pro přípravu dílčích PCR reakcí	Templát pro přípravu dílčích PCR	Primery pro přípravu kazety pomocí fúzní PCR	Primery pro přípravu templátu na sekvenování	Primery pro ověření integrace kazety
				Sekvenace	
pRKL2-17	ORF9_D13_F1 ORF9_D13_R1	DNA viru <i>Vaccinia</i>	ORF9_D13_F1 ORF9_D13_R2	5_RACE_O8_K2 in_Kan_rev1	KanV2F 5RACE_O10_K2
	KL_orf6C_Flag2F ORF9_D13_R2	pRKL1-1		5_RACE_O8_K2 in_Kan_rev1	



**Obrázek 55:** Gen *D13L* viru *Vaccinia* nedokáže zastoupit funkci genu *K2ORF9* pGKL plasmidů *in vivo*. Elektroforetogram přípravy kmene *K. lactis* IFO1267\_pRKL2-17. Rekombinantní pGKL plasmid je označen červeně. M: Lambda DNA/Eco130I (*Sfi*I) Marker, 16. Kultivace jedné nezávislé monoklonie kmene IFO1267\_pRKL2-17 v selektivních podmínkách vedla u všech analyzovaných klonů k ponechání plasmidu pGKL2 po 5. i po 20. pasáži. Plasmid pRKL2-17 tak není schopen v buňkách zastoupit funkci plasmidu pGKL2. Velikost pRKL2-17 je 14455 bp.

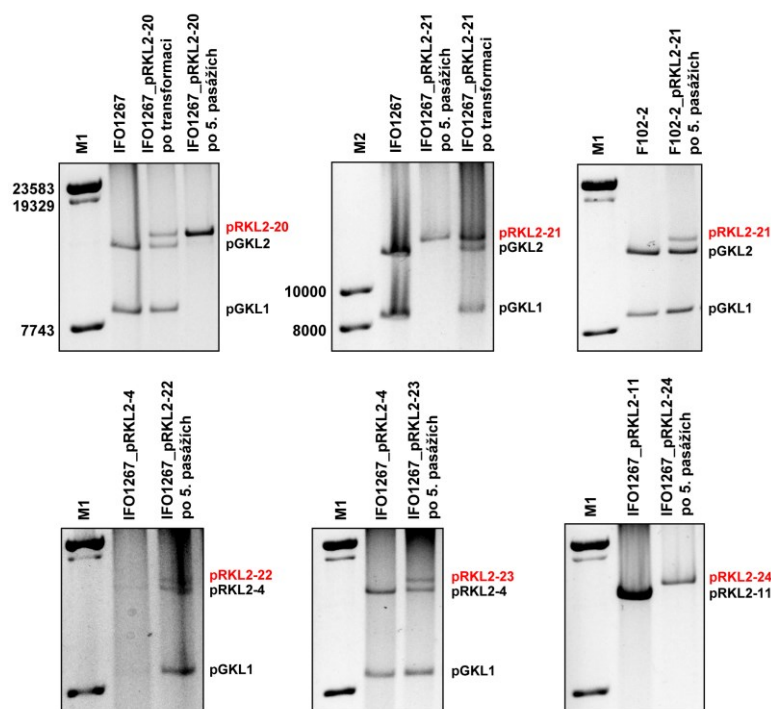
**Závěr:** Plasmid pRKL2-17 není schopen v buňkách zastoupit funkci plasmidu pGKL2. Výsledky naznačují, že gen *D13L* viru *Vaccinia* nedokáže zastoupit funkci esenciálního genu *K2ORF9* pGKL plasmidů *in vivo*.

### 5.1.21 Předpokládaný *scaffold* protein K2ORF9p interaguje s transkripčním aparátem pGKL plasmidů *in vivo*

Vzhledem k tomu, že geny *ORF9* kvasinkových cytoplasmatických lineárních plasmidů kódují předpokládaný *scaffold* protein homologní k proteinu D13 viru *Vaccinia*, je možné, že by kvasinkové lineární plasmidy mohly teoreticky tvořit částice podobné nezralým virovým částicím VACV. Proto jsem se rozhodl detailněji charakterizovat protein K2ORF9p pGKL plasmidů. Především mě zajímalo, zda lze některý konec proteinu K2ORF9p označit proteinovou značkou za současného zachování biologických funkcí tohoto proteinu. Zároveň jsem chtěl ověřit své dřívější výsledky získané proteinovou hmotnostní spektrometrií, které naznačují, že protein K2ORF9p koimunoprecipituje s velkou podjednotkou RNAP a mRNA *capping* enzymem pGKL plasmidů (kapitola 5.1.2.2, Tabulka 14). Předpokládaná interakce K2ORF9p s transkripčním komplexem pGKL plasmidů je zajímavá především z důvodu pravděpodobného shodného poxvirového původu RNA polymerázy, mRNA *capping* enzymu, předpokládané helikázy a předpokládaného *scaffold* proteinu. Výsledkem snažení o charakterizaci K2ORF9p byla příprava 6 kmenů s jednoduše a dvojité rekombinovanými pGKL plasmidy. Postup přípravy jednoduše a dvojité rekombinovaných pGKL plasmidů byl již detailněji popsán v kapitolách 5.1.1.1 a 5.1.2.1. Nejdůležitější informace k přípravě a charakterizaci nových kmenů souvisejících s touto kapitolou jsou shrnuty v Tabulce 32. S přípravou kmenů *K. lactis* IFO1267\_pRKL2-23 a IFO1267\_pRKL2-24 asistovali studenti Přírodovědecké fakulty UK v rámci cvičení Praktikum z molekulární genetiky v roce 2016. Elektroforetogramy detekce

DNA rekombinantních pGKL plasmidů v připravených kmenech jsou na Obrázku 56 a detailní popis charakteristik rekombinantních plasmidů je uveden v Tabulce 6.

Tabulka 32: Primery a templáty pro přípravu a charakterizaci kmenů této kapitoly					
Plasmid	Primery na přípravu kazety pomocí PCR	Templát pro přípravu kazety	Primery na přípravu kazety pomocí fúzní PCR	Primery pro přípravu templátu na sekvenování / Sekvence	Primery pro ověření integrace kazety
pRKL2-20	pGKL_HA-ORF9_F1 pGKL_HA-ORF4_R1	pcDNA™ 5/FRT/TO	pGKL_HA-ORF9_F1 in_ORF9_rev	pGKL_HA-ORF9_F1 in_ORF9_rev	Hygro_rev 5RACE_O7_K2
	pGKL_HA-ORF9_F2 in_ORF9_rev	pGKL2		in_ORF9_rev	
pRKL2-21	pGKL_HA-ORF9z_F1	pcDNA™ 5/FRT/TO	pGKL_HA-ORF9z_F1 in_ORF9_rev	ORF9Rsekven	ORF9Rsekven Hygro_tail
pRKL2-23	pGKL_HA-ORF9z_R1	pGKL2		Hygro_tail	
pRKL2-24	pGKL_HA-ORF9_F2 in_ORF9_rev			Hygro_tail	
pRKL2-22	ORF9-K2_tail_2 pGKL_ORF9-HA_R1	pGKL2	ORF9-K2_tail_2 pGKL_ORF9-HA_R2	ORF9-K2_tail_2 pGKL_ORF9-HA_R2	netestováno
	pGKL_ORF3-HA_R1 pGKL_ORF9-HA_R2	pcDNA™ 5/FRT/TO		ORF9-K2_tail_2	



**Obrázek 56:** Elektroforetogramy připravených kvasinkových kmenů kódujících protein K2ORF9p s HA-značkou. Rekombinantní pGKL plasmidy připravených kmenů jsou označeny červeně. M1: Lambda DNA/Eco130I (*Syl*) Marker, 16; M2: GeneRuler™ 1 kb DNA Ladder. Zobrazeny jsou vybrané monoklonie po pěti pasážích nezávislých klonů získaných transformací a následnou selekcí, případně i nezávislé klony bezprostředně po transformaci. Kultivace kmene *K. lactis* IFO1267\_pRKL2-20 a IFO1267\_pRKL2-21 v selektivních podmínkách vede u všech analyzovaných klonů k vyředění plasmidu pGKL1. V buňkách *S. cerevisiae* F102-2\_pRKL2-21 překvapivě není plasmid pRKL2-21 schopen zastoupit funkci plasmidu pGKL2. Plasmid pRKL2-22 a pRKL2-23 není schopen v buňkách zastoupit funkci nemodifikovaného plasmidu pRKL2-4. S kmeny *K. lactis* IFO1267\_pRKL2-22 a *S. cerevisiae* F102-2\_pRKL2-21 nebylo dále pracováno. Velikost rekombinantních pGKL plasmidů je uvedena v Tabulce 6.

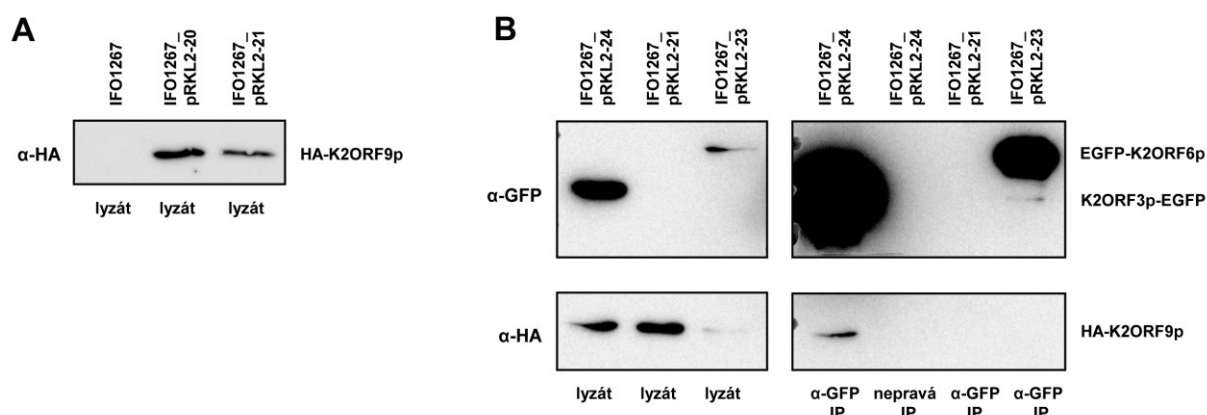
Při přípravách kmene *K. lactis* IFO1267\_pRKL2-20 (HA-K2ORF9p) a obdobného kmene IFO1267\_pRKL2-21 (HA-K2ORF9p), kde je gen *Hyg<sup>R</sup>* orientován v opačné orientaci, byl rekombinantní plasmid schopen zastoupit funkci plasmidu pGKL2, ale vždy zároveň došlo k vyředění plasmidu pGKL1. To může být způsobeno buď jinou mírou produkce rekombinantního proteinu HA-K2ORF9p, která byla pod kontrolou K2UCR4, nebo má rekombinantní protein HA-K2ORF9p mírně odlišné vlastnosti od proteinu K2ORF9p divokého typu, které nestačí v buňce pokrývat požadavky obou plasmidů na tento esenciální protein, což vede k vyředění neautonomního plasmidu pGKL1. Produkce proteinu HA-K2ORF9p ve zmíněných kmenech byla ověřena na úrovni proteinů pomocí metody *western blot* (Obrázek 57A). Pokusil jsem se též připravit plasmid pRKL2-21 v buňkách *S. cerevisiae* F102-2, které jsme původně zamýšleli ve spolupráci s Martinem Fraiberkem (BIOCEV) použít pro kryoimunoelektronovou mikroskopii, kde by vzájemná detekce cytoplasmatické DNA a HA-značeného proteinu mohla vést k detekci případných viru podobných částic. Plasmid pRKL2-21 v buňkách *S. cerevisiae* F102-2 však není soběstačný (Obrázek 56).

Pro validaci předchozích výsledků hmotnostní spektrometrie byly připraveny kmeny kódující vybraný protein transkripčního aparátu pGKL plasmidů ve fúzi s EGFP a předpokládaný *scaffold* protein s HA-značkou. Plasmid pRKL2-22 v buňkách nebyl schopen zastoupit funkci cíleného plasmidu pRKL2-4, a proto s ním nebylo dále pracováno. Překvapivých výsledků bylo získáno při přípravě kmenů *K. lactis* IFO1267\_pRKL2-23 (EGFP-K2ORF6p, HA-K2ORF9p) a IFO1267\_pRKL2-24 (K2ORF3p-EGFP, HA-K2ORF9p), neboť protein HA-K2ORF9p je *in vivo* schopen zastoupit všechny biologické funkce nutné pro udržení plasmidů odvozených z plasmidu pGKL2 u kmenů *K. lactis* IFO1267\_pRKL2-20, IFO1267\_pRKL2-21 a IFO1267\_pRKL2-24, ale překvapivě ne u kmene IFO1267\_pRKL2-23 (Obrázek 56).

Z kmenů IFO1267\_pRKL2-23 a IFO1267\_pRKL2-24 byly připraveny proteinové lyzáty (metoda 4.2.4.4.2) a z proteinových lyzátů byly proteiny EGFP-K2ORF6p a K2ORF3p-EGFP imunoprecipitovány pomocí GFP-Trap<sup>®</sup>\_A nebo nepravě imunoprecipitovány pomocí prázdných agarosových kuliček (metoda 4.2.4.6). Vzorky lyzátů a imunoprecipitovaných proteinů byly rozděleny SDS-PAGE proteinovou elektroforézou (metoda 4.2.4.8), přeneseny na membránu pomocí metody *western blot* (metoda 4.2.4.9) a vybrané proteiny byly chemiluminiscenčně imunodetekovány (metoda 4.2.4.10). Výsledky jasně ukazují, že předpokládaný *scaffold* protein K2ORF9p specificky interaguje s mRNA *capping* enzymem pGKL plasmidů *in vivo* (Obrázek 57B), což potvrzuje dřívejší výsledky hmotnostní spektrometrie, naznačující, že K2ORF9p specificky interaguje s mRNA *capping* enzymem a velkou podjednotkou RNAP pGKL plasmidů (Tabulka 14). Výsledky také překvapivě ukazují,



že značený *scaffold* protein K2ORF9p neinteraguje se značenou velkou podjednotkou RNAP pGKL plasmidů (Obrázek 57B), což je s největší pravděpodobností důvodem, proč je plasmid pRKL2-23 v buňkách kmene *K. lactis* IFO1267\_pRKL2-23 závislý na plasmidu pRKL2-4.



**Obrázek 57:** Předpokládaný *scaffold* protein K2ORF9p interaguje s transkripčním aparátem pGKL plasmidů *in vivo*. (A) *Western blot* analýza lyzátů z kmenů *K. lactis* IFO1267\_pRKL2-20 (HA-K2ORF9p), IFO1267\_pRKL2-21 (HA-K2ORF9p) a kontrolního kmene IFO1267. (B) *Western blot* imunoprecipitovaných ( $\alpha$ -GFP IP) a nepravě imunoprecipitovaných (nepravá IP) proteinů z IFO1267\_pRKL2-23 (EGFP-K2ORF6p, HA-K2ORF9p), IFO1267\_pRKL2-24 (K2ORF3p-EGFP, HA-K2ORF9p) nebo kontrolních buněk IFO1267\_pRKL2-21 (HA-K2ORF9p). Protilátky používané pro imunodetekci jsou uvedeny na levé straně výřezu. Detekované proteiny jsou uvedeny na pravé straně výřezu. Pozice identifikovaných proteinů odpovídaly teoretické molekulové hmotnosti rekombinantních proteinů plné délky, což bylo určeno podle pozic obarvených proteinů PageRuler™ Prestained Protein Ladder (nejsou zobrazeny). Lyzát představoval přibližně 1/100 vzorku a IP představoval přibližně 1/2 vzorku. Nepravá IP byla provedena použitím prázdných agarosových kuliček, pravá  $\alpha$ -GFP IP byla provedena pomocí GFP-Trap®\_A. Pro detekci byly použity následující primární protilátky:  $\alpha$ -HA 6E2 (ředění 1:1000) a  $\alpha$ -GFP B-2 (1:1000). Primární protilátky byly použity v kombinaci se sekundární protiláčkou  $\alpha$ -mouse IgG-HRP (1:5000).

**Závěr:** Výsledky ukazují, že předpokládaný *scaffold* protein K2ORF9p specificky interaguje s transkripčním aparátem pGKL plasmidů *in vivo*. Tato specifická interakce se jeví jako esenciální pro udržení pGKL plasmidů v hostitelských buňkách, protože k ní v buňkách kmene *K. lactis* IFO1267\_pRKL2-23 s nesoběstačným plasmidem pRKL2-23 zřejmě nedochází. Proteinové produkty všech 5 otevřených čtecích rámců pGKL plasmidů s předpokládaným poxvirovým původem interagují *in vivo*. Vzhledem ke kódovací kapacitě 15 ORF pGKL plasmidů je tak velice pravděpodobné, že kvasinkové cytoplasmatické lineární plasmidy vznikly redukcí a domestikací viru blízkého virům z čeledi *Poxviridae*.



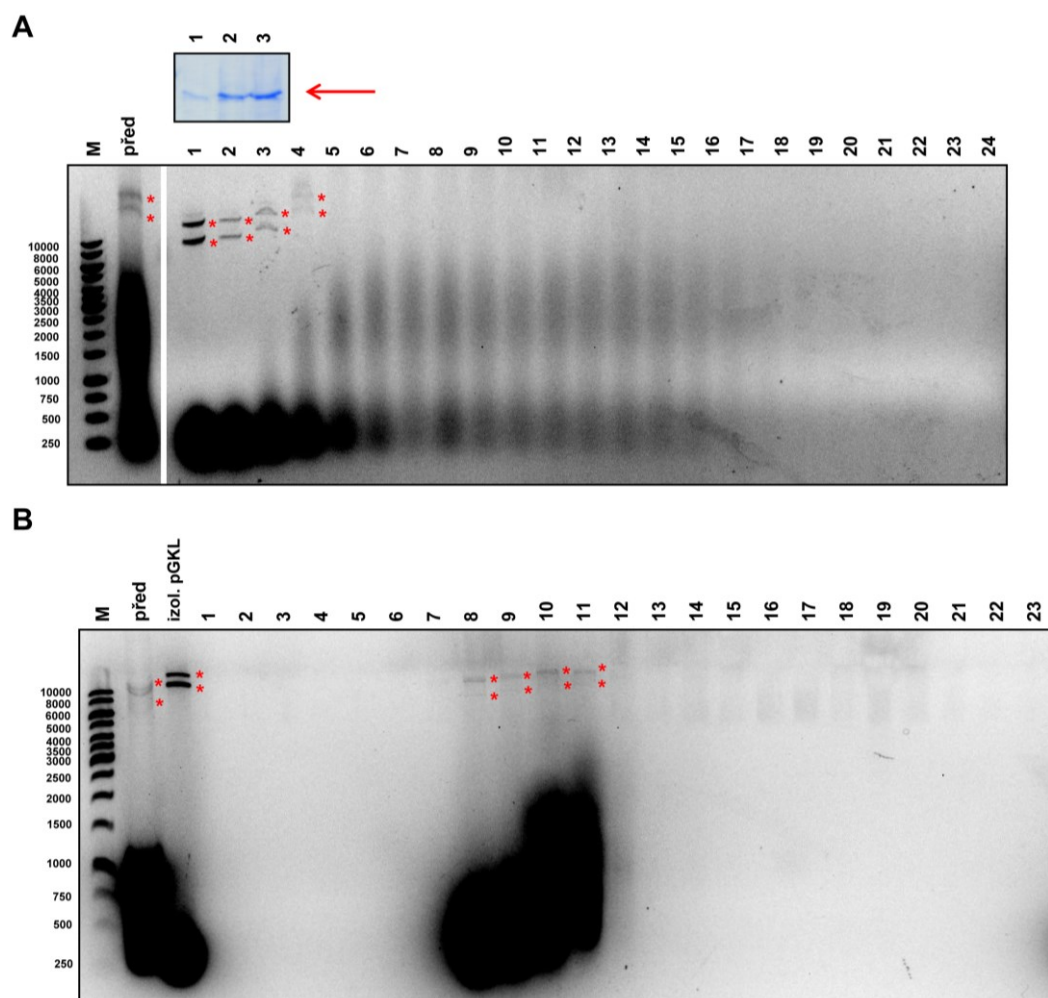
### 5.1.22 Předpokládané viru podobné částice pGKL plasmidů se nepodařilo izolovat ultracentrifugací

Vzhledem k tomu, že by se předpokládaný *scaffold* protein K2ORF9p mohl teoreticky podílet na vytváření vnitrobuněčných membránových struktur kolem genomu pGKL plasmidů, podobných nezralým částicím viru *Vaccinia*, pokusil jsem se předpokládané struktury izolovat pomocí ultracentrifugace. Protože nebylo doposud předpokládáno, že by pGKL plasmidy kódovaly protein příbuzný virovým kapsidovým proteinům, přítomnost možných vnitrobuněčných membránových struktur mohla při výzkumu pGKL plasmidů uniknout pozornosti, také díky využití silných iontových detergentů při izolacích pGKL plasmidů (GUNGE *et al.* 1981; STAM *et al.* 1986).

K izolaci předpokládaných viru podobných částic (VLP) byl použit kmen *K. lactis* IFO1267 a IFO1267\_pGKL<sup>-</sup> (negativní kontrola) za využití dvou různých protokolů. První z nich využíval přípravu lyzátu kvasinkových protoplastů a zakoncentrování pGKL plasmidů do polštáře chloridu cesného ultracentrifugací podle (STAM *et al.* 1986), následované ultracentrifugací v gradientu chloridu cesného (metoda 4.2.5.1.1). Na přípravě vzorků, ultracentrifugaci a rozebírání gradientu na jednotlivé frakce se podílela Jana Váňová. Prostory, část potřebného materiálu a vybavení na ultracentrifugaci a rozebrání gradientu mi byly laskavě poskytnuty Laboratoří virologie Přírodovědecké fakulty UK Hanou Španielovou. Detekce DNA pGKL plasmidů v jednotlivých frakcích využívala agarosovou (metoda 4.2.2.3.1) a SDS-agarosovou (metoda 4.2.2.3.2) elektroforézu. Druhý protokol využíval přípravu buněčného lyzátu skleněnými kuličkami, předpokládané zakoncentrování VLP částic ultracentrifugací přes sacharosový pošťár a ultracentrifugaci v lineárním sacharosovém gradientu (metoda 4.2.5.1.2). Přípravu sacharosového gradientu a rozebrání gradientu na jednotlivé frakce provedl kolega Tomáš Mašek. Detekce DNA pGKL plasmidů v jednotlivých frakcích využívala UV spektrofotometrii a organickou extrakci následovanou ethanolovou precipitací a elektroforetickou analýzou nukleových kyselin jednotlivých frakcí.

Bohužel ani jedním protokolem se nepodařilo detekovat nabohacení DNA pGKL plasmidů ve frakcích gradientu, jenž by napovídalo zakoncentrování předpokládaných VLP. V případě ultracentrifugace v gradientu chloridu cesného se podařilo většinu DNA pGKL plasmidů izolovat oddělenou od proteinů či se sníženou elektroforetickou mobilitou, jež se dá vysvětlit přítomností většího množství kopurifikovaného proteinu o molekulové hmotnosti přibližně 20 kDa (Obrázek 58A), který byl hmotnostní spektrometrií určen jako necharakterizovaný buněčný protein (Tabulka 33), částečně podobný proteinu Gre1p *S. cerevisiae*, což je hydrofilin neznámé funkce. V případě ultracentrifugace v sacharosovém gradientu

se DNA pGKL plasmidů detekovatelná barvením pomocí SYBR<sup>TM</sup> Gold nacházela ve frakcích, v kterých sedimentovala všechna ostatní izolovaná nukleová kyselina o velikostech 1–3000 nukleotidů, takže se s největší pravděpodobností nejednalo o pGKL plasmidy přítomné ve VLP částicích, které by se díky předpokládanému odlišnému sedimentačnímu koeficientu nacházely v jiných frakcích gradientu (Obrázek 58B).



**Obrázek 58:** Předpokládané viru podobné částice pGKL plasmidů se pravděpodobně nepodařilo izolovat ultracentrifugací. **(A)** Elektroforetogramy dokumentující izolaci předpokládaných viru podobných částit pomocí ultracentrifugace v gradientu chloridu cesného z buněk kmene *K. lactis* IFO1267 pomocí metody 4.2.5.1.1. Horní výřez znázorňuje úsek polyakrylamidového gelu obarveného pomocí *Coomassie Blue*, který obsahuje protein (označen červenou šipkou) kopurifikující s DNA pGKL plasmidů ve frakcích 1–3. Spodní výřez znázorňuje analýzu nukleových kyselin pomocí agarosové elektroforézy a barvení pomocí ethidium bromidu. Zobrazen je vzorek před ultracentrifugací (před) a vzorky jednotlivých frakcí po ultracentrifugaci, očíslované vzestupně ode dna zkumavky. Zřetelná DNA pGKL plasmidů je označena červenými hvězdičkami. M: GeneRuler<sup>TM</sup> 1 kb DNA Ladder. **(B)** Elektroforetogramy dokumentující izolaci předpokládaných viru podobných částit pomocí ultracentrifugace v gradientu sacharosy z buněk kmene *K. lactis* IFO1267 pomocí metody 4.2.5.1.2. Výřez znázorňuje analýzu sražených deproteinovaných nukleových kyselin pomocí agarosové elektroforézy a barvení pomocí SYBR<sup>TM</sup> Gold. Zobrazen je vzorek před ultracentrifugací (před), frakce 1 z Obrázku 58A, obsahující čisté a koncentrované pGKL plasmidy (izol. pGKL), a vzorky jednotlivých frakcí po ultracentrifugaci, očíslované vzestupně od vrcholu zkumavky. Zřetelná DNA pGKL plasmidů je označena červenými hvězdičkami.

**Tabulka 33:** Protein kopurifikující s DNA pGKL plasmidů

Protein	Uniprot	kDa	Počet peptidů	Pokrytí
necharakterizovaný protein	Q6CT14	20,1	17	83 %

**Závěr:** Výsledky ukazují, že pomocí ultracentrifugace pravděpodobně není možné izolovat předpokládané viru podobné částice pGKL plasmidů. Pomocí ultracentrifugace v gradientu chloridu cesného se však podařilo oddělit pGKL plasmidy od kontaminujících biopolymerů v zatím nejvyšší popsané čistotě bez použití proteáz a ribonukleáz. Ultracentrifugací v gradientu chloridu cesného s pGKL plasmidy kopurifikuje protein, což by napovídalo možné asociaci tohoto proteinu s plasmidovou DNA *in vivo*; vzhledem k vysoké koncentraci soli ve vzorku se však nejspíše jedná o nefyziologickou interakci. Předpokládané VLP částice pGKL plasmidů by mohly být strukturně podobné nezralým částicím poxvirů, které jsou však při izolacích pomocí ultracentrifugace nestabilní (VANSLYKE *et al.* 1993).

## 5.2 POXVIRY

### 5.2.1 Transkripty viru *Vaccinia* obsahující 5' poly(A) leader často neobsahují 5' mRNA čepičku

Přítomnost 5' čepičkové struktury na transkriptech genů viru *Vaccinia* je známá již velice dlouhou dobu, přičemž VACV byl první DNA virus, u kterého byla tato RNA struktura identifikována (URUSHIBARA *et al.* 1975; WEI a MOSS 1975). Analýzou celkové časné a postreplikativní mRNA VACV syntetizované *in vivo* bylo odhadnuto, že jedna 5' čepičková struktura se vyskytuje na každých 1000 nukleotidů časné mRNA a na každých 1250 nukleotidů postreplikativní mRNA (BOONE a MOSS 1977). Obdobné výsledky byly získány i dřívejší analýzou celkové časné mRNA VACV syntetizované purifikovanými jádry částic VACV *in vitro*, přičemž průměrná délka časné mRNA syntetizované *in vitro* je 1000 nukleotidů, díky čemuž se předpokládalo, že naprostá většina transkriptů VACV bude obsahovat 5' mRNA čepičku *in vivo* (KATES a BEESON 1970a; WEI a MOSS 1974, 1975). Přítomnost 5' čepičkové struktury na úrovni jednotlivých transkriptů však nikdy zkoumána nebyla, přičemž pro celogenomové analýzy 5' konců transkriptů VACV časných i postreplikativních genů byly selektovány pouze transkripty obsahující 5' čepičkovou strukturu (YANG *et al.* 2011a; YANG *et al.* 2012).

Díky výsledkům 5' RACE-PCR analýzy všech genů pGKL plasmidů provedené V. Vopálenským, kde se mi podařilo detekovat signifikantní negativní korelaci mezi počtem netemplátových adenosinů 5' poly(A) leaderu a přítomností 5' mRNA čepičky, a také díky předpokládané tvorbě 5' poly(A) leaderu mRNA pGKL plasmidů obdobným mechanismem jako u mRNA VACV (kapitola 5.1.11) mě začalo zajímat především složení 5' konců transkriptů VACV obsahujících 5' poly(A) leader, syntetizovaných *in vivo*. Proto jsme navázali spolupráci se Zorou Mělkovou (1. Lékařská fakulta Univerzity Karlovy, BIOCEV), která nám dvakrát laskavě poskytla lyzát buněk savčích tkáňových linií infikovaných virem *Vaccinia* kmene Western Reserve. V prvním případě se jednalo o ~8 milionů HeLa buněk infikovaných VACV s multiplicitou infekce 7,5 PFU/buňku, lyzovaných 4 hodiny po infekci (použity pro 5' RACE-PCR analýzu mRNA časných a středních genů VACV), v druhém případě se jednalo o ~1,4 milionů HeLa buněk infikovaných VACV s multiplicitou infekce 3 PFU/buňku, lyzovaných 12 hodin po infekci (použity pro 5' RACE-PCR analýzu mRNA pozdních genů VACV). Kultivace HeLa buněk a jejich infekce VACV byla provedena v laboratoři Zory Mělkové, a proto v této práci nejsou uvedeny příslušné metody a materiál.

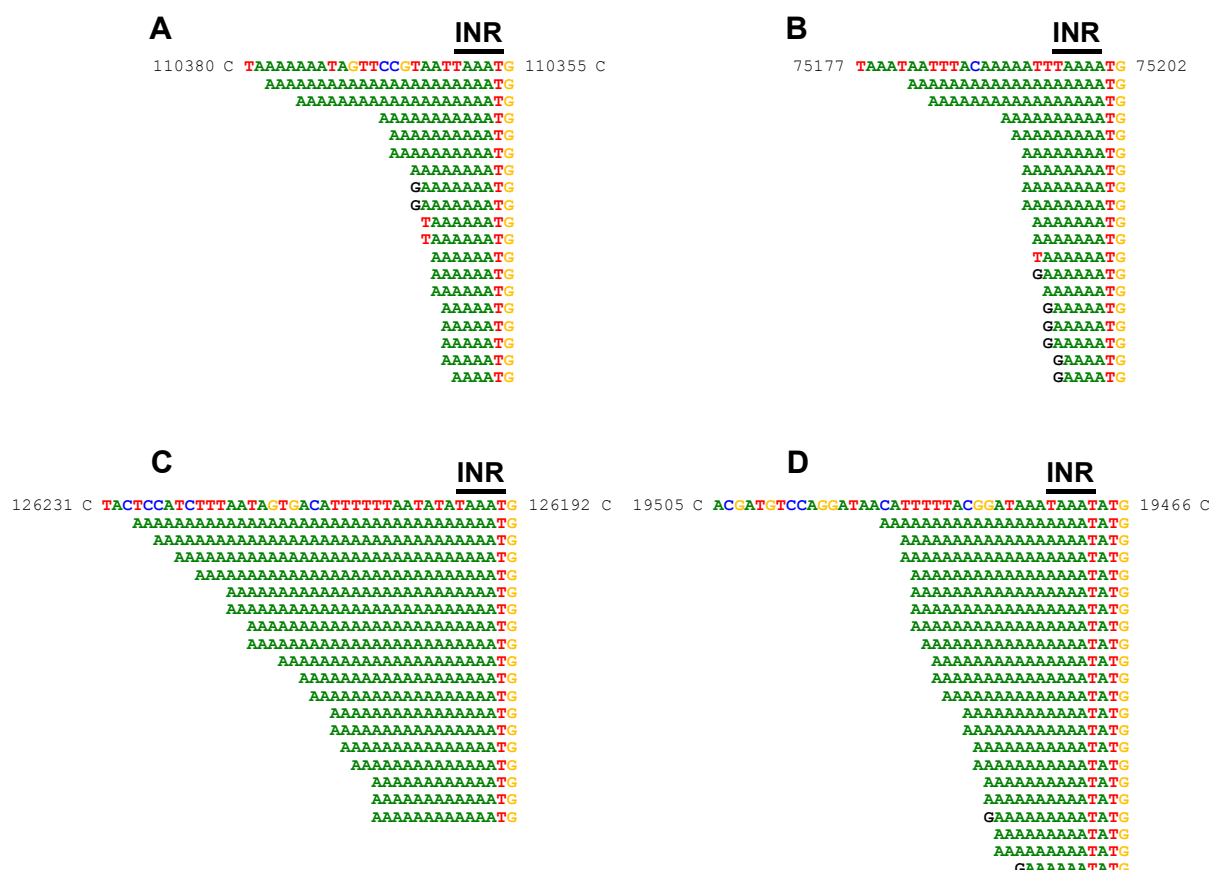
Celková RNA z infikovaných buněk a neinfikovaných kontrolních buněk byla izolována pomocí metody 4.2.3.1.3. Zbývající DNA ve vzorcích byla odstraněna deoxyribonukleázou I

(metoda 4.2.3.1.4) a kvalita RNA byla ověřena pomocí agarosové elektroforézy (metody 4.2.3.3.1 a 4.2.3.3.2) a UV spektrofotometrie (4.2.3.2). Purifikovaná celková RNA byla použita k reverzní transkripci využívající náhodné primery (metody 4.2.3.6.1 a 4.2.3.6.2). Vzniklá cDNA byla purifikována (metoda 4.2.2.1.7) a použita k syntéze homopolymerního 3' konce (metoda 4.2.2.4.6). Vzniklá cDNA s 3' poly(G) koncem byla použita pro 5' RACE-PCR s primery uvedenými v Tabulce 34. Produkty 5' RACE-PCR reakcí byly rozděleny agarosovou elektroforézou (metoda 4.2.2.3.1), specifický produkt byl izolován (metoda 4.2.2.1.5) a klonován do vektoru pCR<sup>TM</sup>4-TOPO<sup>®</sup> (metoda 4.2.2.4.5). Vektory byly transformovány do buněk *E. coli* XL-1 Blue (metoda 4.2.2.5.1). Ze získaných monoklonů byly vektory následně izolovány (metoda 4.2.2.1.1) a charakterizovány pomocí restrikčního štěpení (metoda 4.2.2.4.1) následovaného agarosovou elektroforézou (metoda 4.2.2.3.1). Z kolonií, jejichž vektory vykazovaly přítomnost insertu odpovídajícímu specifickému produktu 5' RACE-PCR reakce, byly vektory izolovány (metoda 4.2.2.1.2), jejich kvalita byla ověřena UV spektrofotometrií (4.2.2.2), a následně byly podrobeny sekvenaci DNA (metody 4.2.2.8.4 a 4.2.2.8.5) v oblasti insertu primery uvedenými v Tabulce 34. Elektroforetogramy detekce produktů 5' RACE-PCR vybraných genů VACV jsou na Obrázku 59. Sekvence jednotlivých cDNA klonů 5' konců mRNA časných genů VACV jsou na Obrázku 60. Sekvence jednotlivých cDNA klonů 5' konců mRNA středních a pozdních genů VACV jsou na Obrázku 61.

Tabulka 34: 5' RACE-PCR analýza transkriptů viru <i>Vaccinia</i>						
Gen	Třída	Primery pro 5' RACE-PCR (35 cyklů)	Produkt	Klonováno do vektoru	Restrikční analýza vektoru	Sekvenace vektoru
A5R	časný	olig2(dC)anchor 5RACE_VV_A5R	~416 bp	pCR <sup>TM</sup> 4-TOPO <sup>®</sup>	<i>EcoRI</i> <i>BclI</i>	T3
D12L	časný	olig2(dC)anchor 5RACE_VV_D12L	~462 bp	pCR <sup>TM</sup> 4-TOPO <sup>®</sup>	<i>EcoRI</i> <i>EcoRV</i>	T7
H5R	časný	olig2(dC)anchor 5RACE_VV_H5R	~479 bp	pCR <sup>TM</sup> 4-TOPO <sup>®</sup>	<i>EcoRI</i> <i>EcoRV</i>	T7
I4L	časný	olig2(dC)anchor 5RACE_VV_I4L	~556 bp	pCR <sup>TM</sup> 4-TOPO <sup>®</sup>	<i>EcoRI</i> <i>NcoI</i>	T7
J6R	časný	olig2(dC)anchor 5RACE_VV_J6R	~521 bp	pCR <sup>TM</sup> 4-TOPO <sup>®</sup>	<i>EcoRI</i> <i>Acc65I</i>	T7
A1L	střední	olig2(dC)anchor 5RACE_VV_A1L	~433 bp	pCR <sup>TM</sup> 4-TOPO <sup>®</sup>	<i>EcoRI</i> <i>Bpu1102I</i>	T7
G8R	střední	olig2(dC)anchor 5RACE_VV_G8R	~493 bp	pCR <sup>TM</sup> 4-TOPO <sup>®</sup>	<i>EcoRI</i> <i>Cfr42I</i>	T3
A17L	pozdní	olig2(dC)anchor 5RACE_VV_A17L	~538 bp	pCR <sup>TM</sup> 4-TOPO <sup>®</sup>	<i>EcoRI</i>	T7
C3L	pozdní	olig2(dC)anchor 5RACE_VV_C3L	~460 bp	pCR <sup>TM</sup> 4-TOPO <sup>®</sup>	<i>EcoRI</i> <i>XhoI</i>	T7

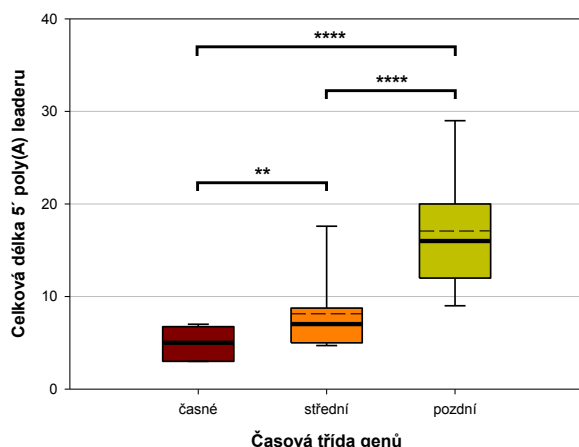






**Obrázek 61:** Transkripty viru *Vaccinia* s 5' poly(A) leaderem často neobsahují 5' mRNA čepičku. (A) 5' RACE-PCR analýza středního genu *A1L*. V tomto a následujících panelech odpovídá horní sekvence templátové (virové) DNA s anotovaným iniciátorovým regionem (INR) podle (YANG *et al.* 2012); sekvence umístěné níže představují jednotlivé sekvenované cDNA klony. 5' nepřekládaná oblast je zobrazena v plném rozsahu až do start kodónu translace, ATG. Guanosinové zbytky odpovídající původním 5' mRNA čepičkám, přítomné v některých klonech cDNA, jsou vyznačeny černě. (B) 5' RACE-PCR analýza středního genu *G8R*. (C) 5' RACE-PCR analýza pozdního genu *A17L*. (D) 5' RACE-PCR analýza pozdního genu *C3L*.

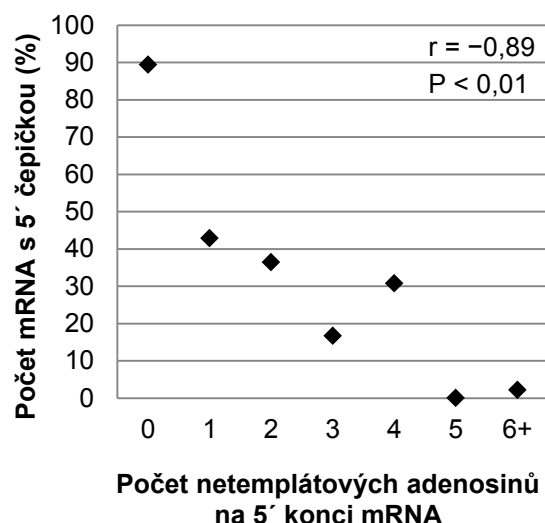
Výsledky překvapivě ukázaly, že zatímco naprostá většina transkriptů časných genů viru *Vaccinia* obsahuje 5' čepičkovou strukturu, tak u transkriptů časných genů s 5' poly(A) leaderem obsahuje 5' čepičkovou strukturu polovina transkriptů a u transkriptů středních a pozdních genů VACV je přítomnost 5' čepičkové struktury mRNA spíše sporadická (Tabulka 35). Shodně s předchozími publikovanými zjištěními se podařilo ověřit, že 5' poly(A) leader mRNA je u každé následující časové třídy genů VACV signifikantně delší (Tabulka 35 a Obrázek 62). Stejně jako u pGKL plasmidů přítomnost 5' čepičky na transkriptech VACV signifikantně negativně koreluje s rostoucím počtem netemplátových adenosinů 5' poly(A) leaderu mRNA (Tabulka 36 a Obrázek 63).



**Tabulka 35:** Četnost 5' mRNA čepičkových struktur a délka 5' poly(A) leaderu transkriptů viru *Vaccinia*

Časová třída	Počet sekvencí	Frekvence výskytu 5' čepičky	Medián (průměr) celkové délky 5' poly(A) leaderu
časně	59	86 %	-
časně s 5' poly(A) leaderem	32	50 %	5 (5)
střední	36	22 %	7 (8)
pozdní	39	5 %	16 (17)

**Obrázek 62:** Vyrůstající délka 5' poly(A) leaderu transkriptů viru *Vaccinia* u následných časových tříd genů. Krabicový graf celkové délky 5' poly(A) leaderu (zahrnuje templátové i netemplátové adenosiny) u transkriptů časných (červeně), středních (oranžově) a pozdních (žlutě) genů viru *Vaccinia*. Spodní a vrchní část boxu indikuje první a třetí kvartil. Úsečky indikují 10. a 90. percentil. Odlehlé hodnoty nejsou vyznačeny. Medián je vyznačen tlustou čarou. Aritmetický průměr je vyznačen přerušovanou čarou. Normální rozdělení dat bylo testováno pomocí Shapiro-Wilk testu. Protože některé kategorie nesplňovaly normální rozdělení, byly data statisticky vyhodnoceny pomocí neparametrického Kruskal-Wallis testu následovaného *post-hoc* Dunn testem s Benjamini-Hochberg FDR korekcí. \*\*: hladina signifikance  $P < 0,01$ . \*\*\*\*: hladina signifikance  $P < 0,0001$ . Celkem bylo k analýze použito 107 sekvencí 5' RACE-PCR klonů genů, jejichž promotory obsahují iniciátorový region. V ojedinělých případech některé sekvenční v oblasti 5' poly(A) leaderu obsahovaly jeden nukleotid jiný, než adenosin. Pro statistické účely bylo s jinými nukleotidy těchto sekvencí zacházeno tak, jako by se jednalo o adenosiny.



**Tabulka 36:** Četnost 5' mRNA čepiček u transkriptů viru *Vaccinia* s různým počtem netemplátových adenosinů 5' poly(A) leaderu

Počet netemplátových nukleotidů	Počet sekvencí	Počet sekvencí s 5' čepičkou	Frekvence výskytu 5' čepičky
0	66	59	89 %
1	14	6	43 %
2	11	4	36 %
3	12	2	17 %
4	13	4	31 %
5	3	0	0 %
6+	47	1	2,1 %

**Obrázek 63:** Signifikantní negativní korelace mezi počtem netemplátových adenosinů 5' poly(A) leaderu a přítomností 5' mRNA čepičky na transkriptech viru *Vaccinia*. Bodový graf znázorňuje frekvenci výskytu 5' mRNA čepičkových struktur u transkriptů jednotlivých tříd. Písmeno *r* značí hodnotu Pearsonova korelačního koeficientu. Výsledky jsou signifikantní na hladině  $P < 0,01$ . Celkem bylo k analýze použito 166 sekvencí cDNA klonů získaných pomocí 5' RACE PCR. V ojedinělých případech některé sekvenční v oblasti netemplátového 5' poly(A) leaderu obsahovaly jeden nukleotid jiný, než adenosin. Pro statistické účely bylo s jinými nukleotidy těchto sekvencí zacházeno tak, jako by se jednalo o adenosiny.

**Závěr:** Výsledky 5' RACE-PCR analýzy vybraných genů překvapivě ukazují, že ne všechny transkripty viru *Vaccinia* obsahují 5' čepičkovou strukturu, přičemž množství detekovaných transkriptů s 5' čepičkovou strukturou u následných časových tříd genů viru *Vaccinia* postupně klesá. Výsledky též potvrdily již dříve popsany signifikantní rozdíl celkové délky 5' poly(A) leaderu u transkriptů různých časových tříd genů viru *Vaccinia*. Podobně jako v případě pGKL plasmidů (Tabulka 24 a Obrázek 44) výsledky poukázaly na signifikantní negativní korelaci mezi počtem netemplátových adenosinů 5' poly(A) leaderu a přítomností 5' mRNA čepičky na transkriptech viru *Vaccinia*. Výsledky nasvědčují, že za rozdílnou míru výskytu 5' čepičkové struktury na transkriptech různých časových tříd genů viru *Vaccinia* by mohla být primárně zodpovědná rozdílná míra 5' polyadenylace těchto transkriptů.

### 5.2.2 Pozdní transkripty viru *Vaccinia* obsahující 5' mRNA čepičku mají kratší 5' poly(A) leader

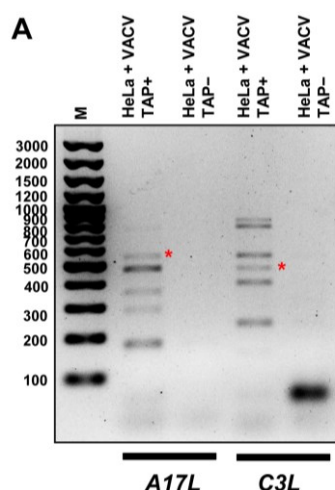
Výsledky získané u pGKL plasmidů i u viru *Vaccinia* shodně poukazují na globální signifikantní negativní korelaci mezi počtem netemplátových adenosinů 5' poly(A) leaderu transkriptu a přítomností 5' mRNA čepičky. Malá četnost výskytu 5' mRNA čepičky na transkriptech s delším 5' poly(A) leaderem znemožňuje kvantifikovat rozdíl v délce transkriptů obsahujících a neobsahujících 5' mRNA čepičku na úrovni jednotlivých genů. Rozhodl jsem se proto kvantifikovat rozdíl v délce transkriptů obsahujících a neobsahujících 5' mRNA čepičku s využitím nabohacení transkriptů daného genu obsahujících 5' mRNA čepičku pomocí soupravy FirstChoice® RLM-RACE Kit (Invitrogen), jejíž chemikálie mi byly laskavě poskytnuty Liborem Krásným. Touto metodou jsem se rozhodl detekovat transkripty pozdních genů viru *Vaccinia*, kde by díky značné délce 5' poly(A) leaderu mRNA těchto genů měl být případný rozdíl v délce transkriptů v závislosti na přítomnosti 5' mRNA čepičky nejvíce markantní.

Stejný vzorek purifikované celkové RNA z infikovaných buněk a neinfikovaných kontrolních buněk jako v předchozí kapitole byl použit k RLM-RACE experimentům (4.2.3.4.1) a následně k reverzní transkripci využívající náhodné primery (metody 4.2.3.6.1 a 4.2.3.6.2). Vzniklá cDNA byla purifikována (metoda 4.2.2.1.7) a použita pro 5' RACE-PCR s primery uvedenými v Tabulce 37. Produkty 5' RACE-PCR reakcí byly rozděleny agarosovou elektroforézou (metoda 4.2.2.3.1), specifický produkt byl izolován (metoda 4.2.2.1.5) a klonován do vektoru pCR™ 4-TOPO® (metoda 4.2.2.4.5). Vektory byly transformovány do buněk *E. coli* XL-1 Blue (metoda 4.2.2.5.1). Ze získaných monoklonií byly vektory následně izolovány (metoda 4.2.2.1.1) a charakterizovány pomocí restriktčního štěpení (metoda 4.2.2.4.1) následovaného agarosovou elektroforézou (metoda 4.2.2.3.1). Z kolonií, jejichž vektory

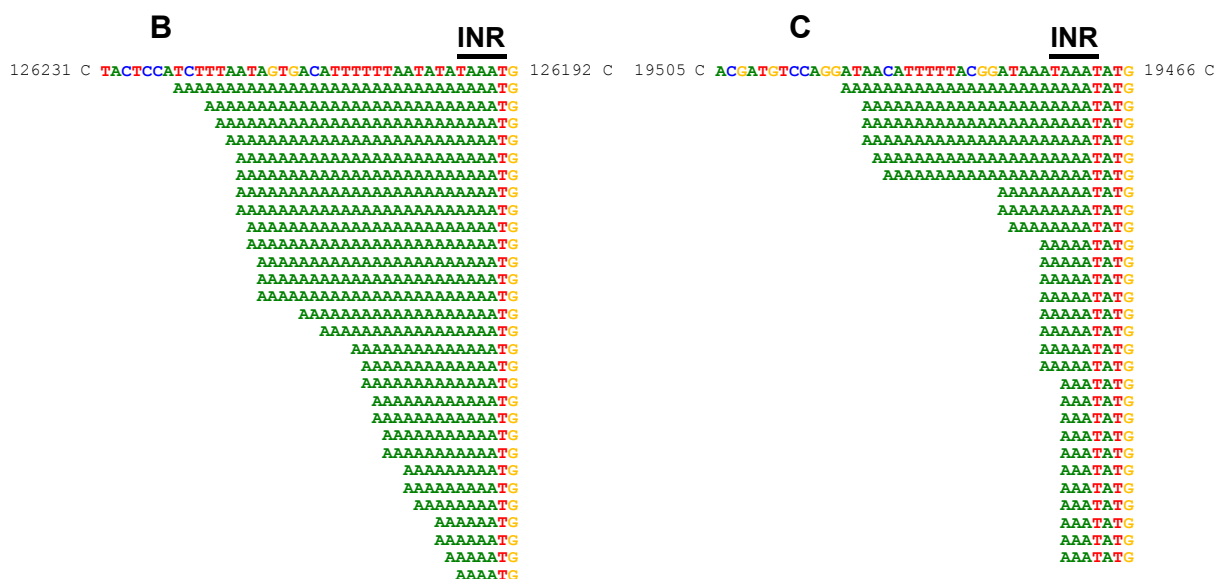
vykazovaly přítomnost insertu odpovídajícímu specifickému produktu 5' RACE-PCR reakce, byly vektory izolovány (metoda 4.2.2.1.2), jejich kvalita byla ověřena UV spektrofotometrií (4.2.2.2), a následně byly podrobeny sekvenaci DNA (metody 4.2.2.8.4 a 4.2.2.8.5) v oblasti insertu primery uvedenými v Tabulce 37. Elektroforetogramy detekce produktů 5' RACE-PCR vybraných genů VACV jsou na Obrázku 64A. Sekvence jednotlivých cDNA klonů 5' konců mRNA pozdních genů VACV jsou na Obrázku 64B a C.

**Tabulka 37:** 5' RLM-RACE analýza transkriptů viru *Vaccinia*

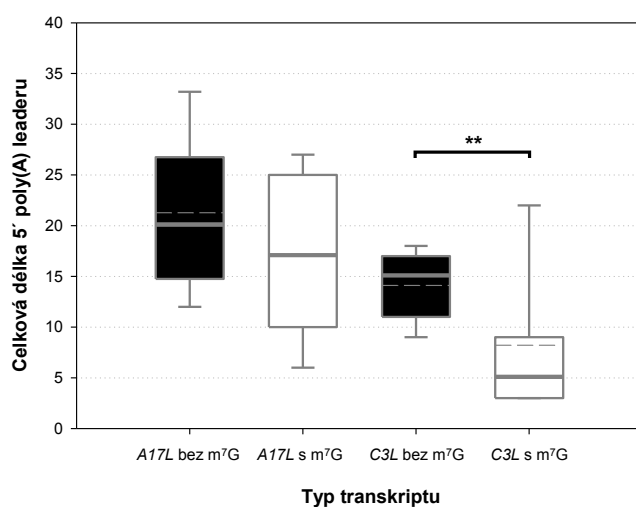
Gen	Třída	Primery pro 5' RACE-PCR (35 cyklů)	Produkt	Klonováno do vektoru	Restrikční analýza vektoru	Sekvenace vektoru
A17L	pozdní	5' RACE Outer Primer 5RACE_VV_A17L	~555 bp	pCR <sup>TM</sup> 4-TOPO®	<i>EcoRI</i>	T7
C3L	pozdní	5' RACE Outer Primer 5RACE_VV_C3L	~475 bp	pCR <sup>TM</sup> 4-TOPO®	<i>EcoRI</i> <i>XhoI</i>	T7



**Obrázek 64:** Pozdní transkripty viru *Vaccinia* obsahující 5' mRNA čepičku mají kratší 5' poly(A) leader. **(A)** Elektroforetogram 5' RLM-RACE analýzy transkriptů genů *A17L* a *C3L* viru *Vaccinia* z HeLa buněk lyzovaných 12 hodin po infekci. Reakce, kdy k RNA byla (TAP+) a nebyla (TAP-) přidána kyselá pyrofosfatáza tabáku jsou označeny. Použité primery a velikost specifických produktů uvádí Tabulka 37. Specifické produkty, které byly klonovány do vektoru pCR<sup>TM</sup>4-TOPO® a použity pro sekvenování, jsou označeny červenými hvězdičkami. Podle restriční analýzy jsou i ostatní produkty specifické. M: GeneRuler<sup>TM</sup> 100 bp Plus DNA Ladder. **(B)** 5' RLM-RACE analýza genu *A17L* viru *Vaccinia*. V tomto a následujících panelech odpovídá horní sekvence templátové (virové) DNA s anotovaným iniciátorovým regionem (INR) podle (YANG *et al.* 2012); sekvence umístěné níže představují jednotlivé sekvenované cDNA klony. 5' nepřekládaná oblast je zobrazena v plném rozsahu až do start kodónu translace, ATG. **(C)** 5' RLM-RACE analýza genu *C3L* viru *Vaccinia*.



Výsledky 5' RLM-RACE v kombinaci s předchozími výsledky klasického 5' RACE-PCR ukázaly, že transkripty obou vybraných pozdních genů viru *Vaccinia* obsahující 5' mRNA čepičkovou strukturu mají kratší 5' poly(A) leader, než transkripty stejných genů neobsahující 5' mRNA čepičku, přičemž v případě genu *C3L* je rozdíl statisticky signifikantní (Tabulka 38 a Obrázek 65). Výsledky 5' RLM-RACE experimentů, detekující více specifických PCR produktů a tedy zřejmě i transkripčních variant obou genů (Obrázek 64A), též poukázaly na známý jev takzvané všudypřítomné iniciace transkripce, kdy k iniciaci transkripce u VACV často dochází i uvnitř kódujících sekvencí otevřených čtecích rámců, což přispívá ke značné transkripční komplexitě VACV (YANG *et al.* 2012).



**Tabulka 38:** Celková délka 5'poly(A) leaderu mRNA pozdních genů viru *Vaccinia* v závislosti na přítomnosti 5' mRNA čepičky

Typ transkriptu	Počet sekvencí	Medián délky	Průměr délky
A17L bez m <sup>7</sup> G	18	20	21,3
A17L s m <sup>7</sup> G	29	17	17,2
C3L bez m <sup>7</sup> G	19	15	14,1
C3L s m <sup>7</sup> G	28	5	8,2

**Obrázek 65:** Pozdní transkripty viru *Vaccinia* obsahující 5' mRNA čepičku mají kratší 5' poly(A) leader. Krabicový graf celkové délky 5' poly(A) leaderu (zahrnuje templátové i netemplátové adenosiny) u transkriptů pozdních genů viru *Vaccinia* neobsahujících (černě) a obsahujících (bíle) 5' mRNA čepičku (m<sup>7</sup>G). Spodní a vrchní část boxu indikuje první a třetí kvartil. Úsečky indikují 10. a 90. percentil. Odlehle hodnoty nejsou vyznačeny. Medián je vyznačen tlustou čarou. Aritmetický průměr je vyznačen přerušovanou čarou. Normální rozdělení dat bylo testováno pomocí Shapiro-Wilk testu. Protože některé kategorie nesplňovaly normální rozdělení, byly data statisticky vyhodnoceny pomocí neparametrického Mann-Whitney testu. \*\*: hladina signifikance  $P < 0,01$ . Celkem bylo k analýze použito 94 sekvencí 5' RACE-PCR a 5' RLM-RACE klonů.

**Závěr:** Výsledky ukazují, že na úrovni transkriptů jednotlivých genů viru *Vaccinia* jsou transkripty obsahující 5' čepičkovou strukturu signifikantně kratší, než transkripty neobsahující 5' čepičkovou strukturu. Výsledky 5' RACE-PCR a 5' RLM-RACE jsou porovnatelné, protože získané sekvence pochází ze stejného výchozího vzorku RNA a v obou případech docházelo k amplifikaci obdobně dlouhých úseků cDNA (rozdíl v délce 17 nt) za obdobných podmínek (35 cyklů).

### 5.2.3 3' konce mRNA postreplikativních genů viru *Vaccinia* se dají mapovat

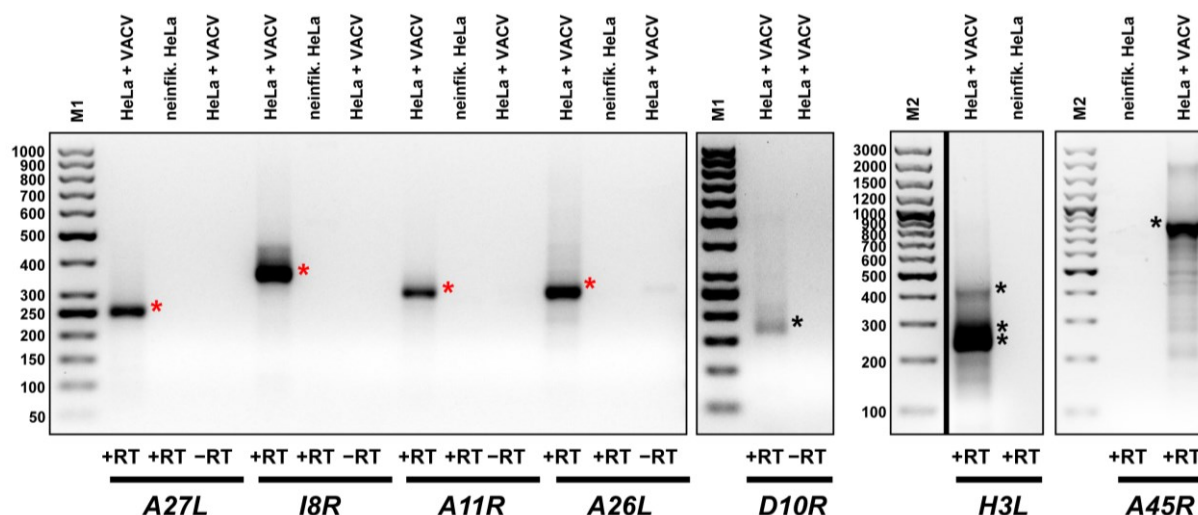
Jak již bylo zmíněno v kapitole 2.2.5.2, tvorba 3' konců transkriptů postreplikativních genů VACV není dopodrobna prozkoumána hlavně díky velké heterogenitě délky 3' konců jednotlivých transkriptů daného genu, což znemožňuje jejich mapování pomocí S1 endonukleázy a detekci délky transkriptů metodou *northern blot* (COOPER *et al.* 1981; MAHR a ROBERTS 1984a; WEIR a MOSS 1984; AHN *et al.* 1992; XIANG *et al.* 1998). Definované 3' konce postreplikativních transkriptů VACV byly zatím popsány jen u tří genů (PATEL a PICKUP 1989; AMEGADZIE *et al.* 1992b; D'COSTA *et al.* 2004). Pomocí celogenomových analýz 5' a 3' konců postreplikativních transkriptů VACV byla kromě výše zmíněné všudypřítomné iniciace transkripce detekována i všudypřítomná tvorba 3' konců mRNA, neboť celkově bylo po 4 a 8 hodinách po infekci buněk virem *Vaccinia* detekováno ~7 (po 4 hod.) a 17 (po 8 hod.) tisíc terminálních nukleotidů mRNA (YANG *et al.* 2012), což je při velikosti genomu VACV 200 tisíc nukleotidů značné množství. Terminální nukleotidy 3' konců mRNA jsou často ve shlučích a ve stejné transkripční orientaci jako přilehlá místa iniciace transkripce, přičemž ~25 koncových nt bývá bohatých na pyrimidinové nukleotidy (YANG *et al.* 2012). Kromě těchto zjištění však celogenomové analýzy transkriptů VACV nepomohly objasnit mechanismus tvorby 3' konců mRNA postreplikativních genů. Díky úspěchům, které jsme v laboratoři zaznamenali použitím metody 3' RACE-PCR pro charakterizaci 3' konců transkriptů všech genů pGKL plasmidů, které se předtím také nedařilo zmapovat pomocí S1 endonukleázy, jsem se rozhodl touto metodikou zkusit detekovat 3' konce transkriptů vybraných postreplikativních genů VACV. Použití metodiky RACE-PCR na modelovém poxviru VACV zatím nebylo publikováno, ačkoliv na jiném poxviru se pomocí této metodiky podařilo charakterizovat 5' a 3' konce mRNA dvou časných genů (BUGERT *et al.* 2000).

Purifikovaná celková RNA byla použita k reverzní transkripci využívající oligod(T)anchor primer (metody 4.2.3.6.1 a 4.2.3.6.3). Vzniklá cDNA byla purifikována (metoda 4.2.2.1.7) a použita pro 3' RACE-PCR s primery uvedenými v Tabulce 39. Produkty 3' RACE-PCR reakcí byly rozděleny agarosovou elektroforézou (metoda 4.2.2.3.1), specifické produkty o vhodné velikosti pro klonování (200–600 bp) byly izolovány (metoda 4.2.2.1.5) a klonovány do vektoru pCR<sup>™</sup>4-TOPO<sup>®</sup> (metoda 4.2.2.4.5). Vektory byly transformovány do buněk *E. coli* XL-1 Blue (metoda 4.2.2.5.1). Ze získaných monoklonií byly vektory následně izolovány (metoda 4.2.2.1.1) a charakterizovány pomocí restriční štěpení enzymy, uvedenými v Tabulce 39. Z kolonií, jejichž vektory vykazovaly přítomnost insertu odpovídajícímu specifickému produktu 3' RACE-PCR reakce, byly vektory izolovány (metoda 4.2.2.1.2), jejich kvalita byla ověřena UV spektrofotometrií (4.2.2.2), a následně byly podrobeny sekvenaci DNA



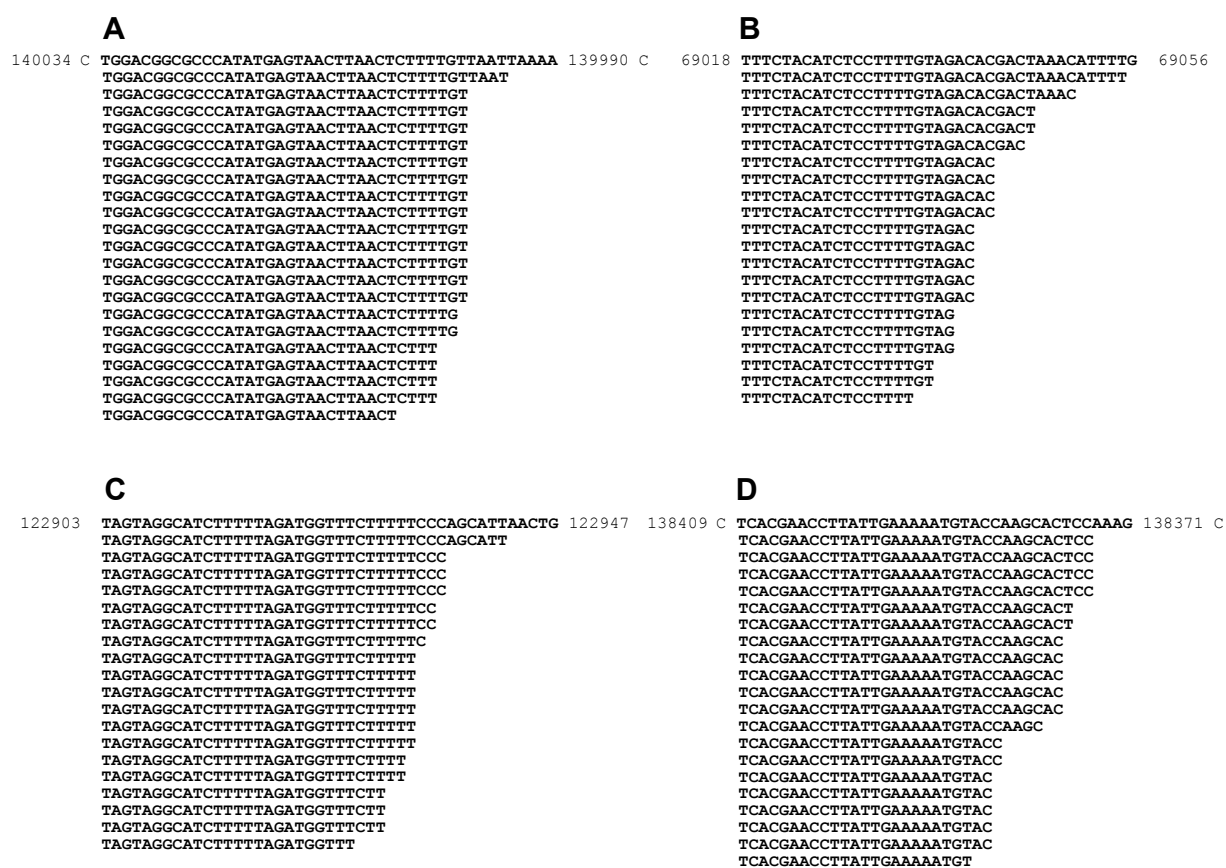
(metoda 4.2.2.8.4) v oblasti insertu primerem T7. Elektroforetogramy detekce některých produktů 3' RACE-PCR jsou na Obrázku 66.

Tabulka 39: 3' RACE-PCR analýza transkriptů postreplikativních genů viru <i>Vaccinia</i>						
Gen	Kategorie	Primery pro 3' RACE-PCR (35 cyklů)	Produkt	Restrikční analýza produktu	Klonováno do vektoru	Restrikční analýza vektoru
A27L	střední	3RACE_VV_A27L PCR-anchor	~253 bp	<i>EcoRV</i>	pCR™4-TOPO®	<i>EcoRI</i> <i>EcoRV</i>
D6R	střední	3RACE_VV_D6R PCR-anchor	nedetekován	–	–	–
D10R	střední	3RACE_VV_D10R PCR-anchor	~175 bp	neprovedena	neklonováno	–
H3L	střední	3RACE_VV_H3L PCR-anchor	více produktů	<i>Mbol</i>	neklonováno	–
I5L	střední	3RACE_VV_I5L PCR-anchor	~150 bp	<i>VspI</i>	neklonováno	–
I8R	střední	3RACE_VV_I8R PCR-anchor	~377 bp	<i>Eco105I</i>	pCR™4-TOPO®	<i>EcoRI</i> <i>Eco105I</i>
O2L	střední	3RACE_VV_O2L PCR-anchor	nedetekován	–	–	–
A11R	pozdní	3RACE_VV_A11R PCR-anchor	~315 bp	<i>Mbol</i>	pCR™4-TOPO®	<i>EcoRI</i> <i>SspI</i>
A21L	pozdní	3RACE_VV_A21L PCR-anchor	nedetekován	–	–	–
A26L	pozdní	3RACE_VV_A26L PCR-anchor	~317 bp	<i>NcoI</i>	pCR™4-TOPO®	<i>EcoRI</i> <i>NcoI</i>
A28L	pozdní	3RACE_VV_A28L PCR-anchor	více produktů	<i>Mbol</i>	neklonováno	–
A45R	pozdní	3RACE_VV_A45R PCR-anchor	~900 bp	<i>Mbol</i>	neklonováno	–



**Obrázek 66:** Elektroforetogramy 3' RACE-PCR analýzy vybraných postreplikativních genů viru *Vaccinia*. Použité primery a velikost specifických produktů je uvedena v Tabulce 39. Specifické produkty, které byly klonovány do vektoru pCR™4-TOPO® a použity pro sekvenování, jsou označeny červenými hvězdičkami. Specifické produkty, které nebyly klonovány, jsou označeny černými hvězdičkami. U genů *D10R* a *H3L* jsou specifické produkty odvozeny z kratších transkriptů neobsahujících stop kodón daného ORF. M1: GeneRuler™ 50 bp Plus DNA Ladder. M2: GeneRuler™ 100 bp Plus DNA Ladder.

K analýze byla použita celková RNA HeLa buněk infikovaných VACV a lyzovaných 4 hodiny po infekci. Některé specifické produkty 3' RACE-PCR pozdních genů VACV se podařilo detekovat i v lyzátu HeLa buněk 12 hodin po infekci VACV (data nezobrazena). Přibližně u poloviny vybraných postreplikativních genů VACV se bez jakékoliv optimalizace podařilo pomocí 3' RACE-PCR detekovat jediný specifický produkt (Tabulka 39). U další poloviny vybraných genů se buď nepodařilo detekovat specifický produkt, nebo bylo detekováno více specifických produktů, což by mohlo být v souladu s obecným předpokladem, že většina transkriptů postreplikativních genů VACV je heterogenní v délce 3' konců (CONDIT a NILES 2002). Pro analýzu sekvencí 3' konců mRNA postreplikativních genů VACV byly vybrány pouze ty geny, kde se pomocí 3' RACE-PCR podařilo detekovat pouze jediný specifický produkt, jenž podle své délky pravděpodobně obsahuje stop kodón daného ORF. Tudíž byly analyzovány 3' konce jen potenciálně translatovatelných transkriptů daných genů, jenž mohou dát vznik proteinovým produktům plné délky. Sekvence jednotlivých cDNA klonů 3' konců mRNA středních genů *A27L* a *I8R* a pozdních genů *A11R* a *A26L* VACV jsou na Obrázku 67.



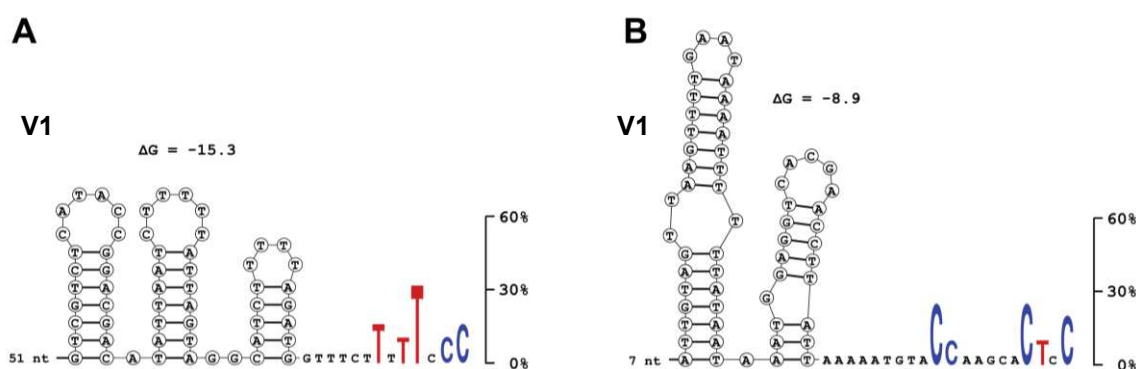
**Obrázek 67:** 3' konce transkriptů postreplikativních genů viru *Vaccinia* se dají mapovat pomocí 3' RACE-PCR analýzy. (A) 3' RACE-PCR analýza středního genu *A27L*. V tomto a následujících panelech odpovídá horní sekvence templátové (virové) DNA; sekvence umístěné níže představují jednotlivé sekvenované cDNA klony (zobrazen je pouze konec 3' nepřekládané oblasti). (B) 3' RACE-PCR analýza středního genu *I8R*. (C) 3' RACE-PCR analýza pozdního genu *A11R*. (D) 3' RACE-PCR analýza pozdního genu *A26L*.

### 5.2.4 3' konce mRNA postreplikativních genů viru *Vaccinia* se nachází poblíž předpokládaných RNA vlásenek

Stejně jako v případě pGKL plasmidů jsem se pokusil bioinformatickými metodami najít sekvenční či strukturní element v blízkosti detekovaných 3' konců mRNA, který by se mohl podílet na formování 3' konců postreplikativních transkriptů VACV. Stejně jako v případě pGKL plasmidů se mi díky *in silico* analýze sekundárních struktur RNA pomocí serveru RNAstructure podařilo identifikovat předpokládané struktury RNA vlásenek poblíž experimentálně určených 3' konců mRNA všech čtyř charakterizovaných postreplikativních genů VACV (Tabulka 40). Předpokládané RNA vlásenky byly relativně blízko za příslušným stop kodómem daného otevřeného čtecího rámce (medián vzdálenosti 13,5 nt) a jejich konce jsou v blízkosti detekovaných 3' terminálních nukleotidů mRNA (medián vzdálenosti 9 nt). Gibbsova volná energie předpokládaných RNA vlásenek je však relativně vysoká (medián  $-7,9$  kcal/mol). Zdá se, že přítomnost jedné či dvou překryvných předpokládaných RNA vlásenkových struktur koresponduje s poměrně jednoduchým vzorem detekovaných 3' konců mRNA čtyř analyzovaných postreplikativních genů VACV (Obrázek 67 a Obrázek 68).

**Tabulka 40:** Předpokládané RNA vlásenky poblíž 3' konců mRNA postreplikativních genů viru *Vaccinia*

Gen	Vlásenka	Vzdálenost k stop kodómu	$\Delta G$	Nukleotidová sekvence 5' → 3'
A27L	V1	-28	-4,7	TGTTCAGACTGGACGGCGCCCATATGAGTAACCTTAACCTC
	V2	-39	-6,2	AAAAGATTGATGTTTCAGACTGGACGGCGCCCATATGAGTAACCTTAACCTTTT
I8R	V1	20	-6,9	CATCTTTTTTTAGAGATGAAAAATTTTCTACATCTCCTTTTGTA
	V2	20	-11,5	CATCTTTTTTTAGAGATGAAAAATTTTCTACATCTCCTTTTGTA
A11R	V1	51	-15,3	GTCGTCTCATACCGGACGACATATAATCTTTTATTAGTAGGCATCTTTTAGATG
A26L	V1	7	-8,9	ATTGTAGTTAAGTTTTGAATAAAATTTTATAATAAATGGAGGTCACGAACCTTATT



**Obrázek 68:** 3' konce transkriptů postreplikativních genů viru *Vaccinia* se nachází poblíž předpokládaných struktur RNA vlásenek. RNA vlásenky jsou zobrazeny jako písmena cDNA nukleotidů v kruzích a pro každou strukturu jsou zobrazeny hodnoty Gibbsovy volné energie ( $\Delta G$ ) v kcal/mol. Vzdálenost začátku RNA vlásenek od stop kodómu genu příslušného transkriptu je uvedena jako počet nukleotidů (nt). Terminální nukleotidy experimentálně stanovených 3' konců cDNA jsou znázorněny jako barevná písmena zvětšená proporcionálně k jejich výskytu (v procentech) v sekvenovaných klonech, když byl detekován stejný terminální nukleotid u sekvencí alespoň dvou nezávislých klonů. (A) Analýza pozdního genu *A11R*. (B) Analýza pozdního genu *A26L*.

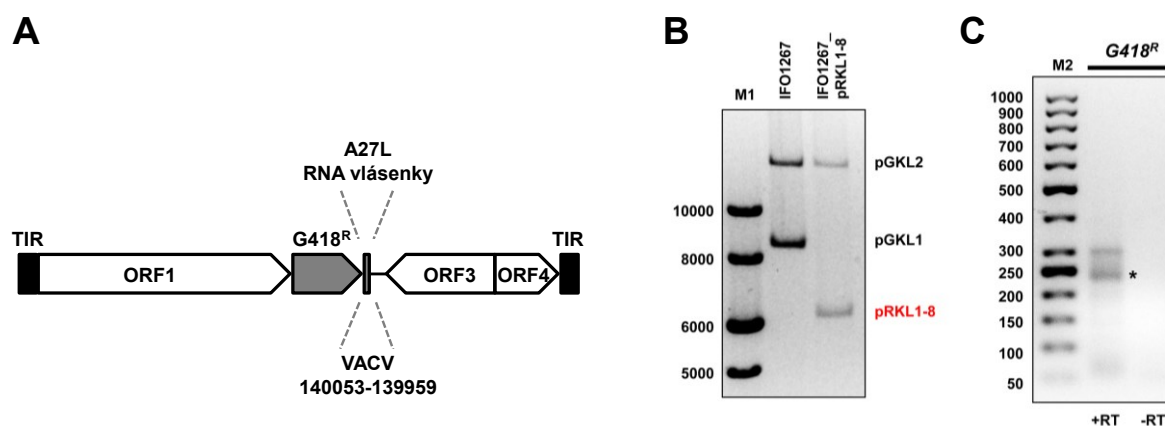
**Závěr:** Výsledky ukazují, že 3' konce transkriptů postreplikativních genů viru *Vaccinia* se nachází poblíž předpokládaných struktur slabých RNA vlásenek. Tyto struktury by se tak možná stejně jako v případě příbuzného transkripčního aparátu pGKL mohly podílet na definované tvorbě 3' konců mRNA postreplikativních genů viru *Vaccinia in vivo*. Předpokládaná vysoká frekvence náhodného výskytu slabých struktur RNA vlásenek by mohla korespondovat s doloženou značnou heterogenitou délky 3' konců mRNA postreplikativních genů viru *Vaccinia*, která znesnadňuje výzkum mechanismů terminace transkripce těchto genů.

### 5.2.5 Předpokládané RNA vlásenky postreplikativních genů viru *Vaccinia* jsou schopny indukovat definovanou tvorbu 3' konců mRNA pGKL plasmidů *in vivo*

Rozhodl jsem se otestovat, zda předpokládané struktury RNA vlásenek dokáží ovlivnit tvorbu 3' konců mRNA *in vivo*. Jelikož veškeré způsoby, jakými by šlo tento jev testovat s využitím VACV, by vyžadovaly poměrně časově náročnou spolupráci s potřebně vybavenou virologickou laboratoří, rozhodl jsem se testovat vliv předpokládaných RNA vlásenek VACV na tvorbu 3' konců mRNA vnesením jejich kódujících sekvencí do pGKL plasmidů, kde se mi již dříve podařilo ve fylogeneticky příbuzném transkripčním aparátu prokázat vliv slabých RNA vlásenek na definovanou tvorbu 3' konců mRNA (kapitola 5.1.17).

Byl připraven kmen *K. lactis* s modifikovaným pGKL plasmidem postupem již dříve detailněji popsáním v kapitole 5.1.1.1. Připravený kmen obsahoval modifikovaný plasmid pGKL1 kódující gen *G418<sup>R</sup>* pod kontrolou promotoru K1UCR2 (gen *G418<sup>R</sup>* vložený místo genu *K1ORF2*), který byl následován oblastí středního genu *A27L* viru *Vaccinia*, která kóduje dvě částečně překryvné předpokládané RNA vlásenky V1 a V2 (IFO1267\_pRKL1-8). Nejdůležitější informace k přípravě a charakterizaci nového kmene jsou shrnuty v Tabulce 41. Schéma kódující kapacity rekombinantního pGKL plasmidu je na Obrázku 69A. Elektroforetogram detekce rekombinantního pGKL plasmidu v připraveném kmeni je na Obrázku 69B.

Tabulka 41: Primery a templáty pro přípravu a charakterizaci kmene této kapitoly						
Plasmid	Primery na přípravu kazety pomocí PCR	Templát pro přípravu kazety	Primery pro ověření integrace kazety a přípravu templátu na sekvenování / Sekvenace	Primery pro 3' RACE-PCR (35 cyklů)	Produkt	Restrikční analýza vektoru
pRKL1-8	pGKL-kanF K1_VV_A27L-term	pRKL1-1	KanVerF ORF3-pGKL1_tail KanVerF	KanVerF anch2	~238 bp	<i>EcoRI</i> <i>NdeI</i>



**Obrázek 69:** Předpokládané RNA vlásenky postreplikativních genů viru *Vaccinia* jsou schopny indukovat definovanou tvorbu 3' konců mRNA pGKL plasmidů *in vivo*. (A) Schematické znázornění rekombinantního plasmidu pRK1-8, kde je gen *G418<sup>R</sup>* následován kódující sekvencí dvou předpokládaných RNA vlásenek transkriptů středního genu *A27L* viru *Vaccinia*. TIR: terminální invertované repetice. (B) Elektroforetogram DNA připraveného kmene *K. lactis* IFO1267\_pRK1-8. Zobrazena je vybraná monoklonie po pěti pasážích nezávislého klonu získaného po transformaci. Rekombinantní pGKL plasmid je označen červeně. Velikost plasmidu pRK1-8 je 6484 bp. M1: GeneRuler™ 1 kb DNA Ladder. (C) Elektroforetogram 3' RACE-PCR analýzy transkriptů genu *G418<sup>R</sup>*. Reverzní transkripce byla prováděna za přítomnosti (+RT) a nepřítomnosti (–RT) reverzní transkriptázy. Pro 3' RACE-PCR byly použity primery uvedené v Tabulce 41. Specifický produkt, který byl klonován do vektoru pCR™4-TOPO® a použit pro sekvenování, je označen hvězdičkou. M2: GeneRuler™ 100 bp Plus DNA Ladder.

Izolace, purifikace a modifikace RNA, stejně jako reverzní transkripce a purifikace cDNA byla provedena stejně, jako je popsáno v kapitole 5.1.17. Purifikovaná cDNA byla použita pro 3' RACE-PCR s primery KanVerF a anch2. Produkty 3' RACE-PCR reakcí byly rozděleny agarosovou elektroforézou (metoda 4.2.2.3.1), specifické produkty byly izolovány (metoda 4.2.2.1.5) a klonovány do vektoru pCR™4-TOPO® (metoda 4.2.2.4.5). Izolace, charakterizace a sekvenování vektorů proběhlo stejně, jako je popsáno v kapitole 5.1.17. Elektroforetogram detekce produktu 3' RACE-PCR je na Obrázku 69C. Výsledky 3' RACE-PCR experimentů z kmene *K. lactis* IFO1267\_pRK1-8 ukázaly, že definované 3' konce mRNA genu *G418<sup>R</sup>* přepisovaného z pGKL plasmidu se dokáží tvořit v blízkosti předpokládaných RNA vlásenkových struktur transkriptů středního genu *A27L* viru *Vaccinia* (Obrázek 70). Umístění terminálních nukleotidů charakterizovaných 3' konců mRNA genů *A27L* a *G418<sup>R</sup>* přepisovaných z genomu VACV a pGKL plasmidu je velmi podobné (Obrázek 70).





## 6 DISKUZE

### 6.1 KVASINKOVÉ CYTOPLASMATICKÉ LINEÁRNÍ PLASMIDY

#### 6.1.1 Integrita aktivního místa methyltransferázy mRNA *capping* enzymu je nezbytná pro udržení pGKL plasmidů *in vivo*

Gen *K2ORF3* kódující mRNA *capping* enzym je esenciální pro udržení pGKL plasmidů v hostitelských buňkách (LARSEN *et al.* 1998). Předpokládaná methyltransferázová doména K2ORF3p byla identifikována na základě bioinformatických analýz detekujících přítomnost konzervovaných motivů I a Ia v kódujících sekvencích genů *ORF3* plasmidů pGKL2 a pSKL (WANG a SHUMAN 1997; LARSEN *et al.* 1998). Methyltransferázovou aktivitu proteinu K2ORF3p se však navzdory několika nezávislým pokusům nepodařilo *in vitro* prokázat (TIGGEMANN *et al.* 2001; VOPÁLENSKÝ 2007). To vyvolalo otázku, zda je v proteinu K2ORF3p lokalizována celá methyltransferázová doména nebo pouze její N-koncová část. Bioinformatickou analýzou se mi podařilo prokázat, že kromě již anotované podobnosti proteinů ORF3p kvasinkových lineárních plasmidů k zhruba prvním dvěma třetinám methyltransferáz vykazují proteiny ORF3p podobnost i k části zbýbající třetiny methyltransferáz (SÝKORA 2013). Model terciární struktury K2ORF3p naznačuje, že methyltransferázová doména je sice částečně redukována, nicméně však obsahuje strukturu sedmi  $\beta$ -listů, charakteristických pro methyltransferázy závislé na SAM (Obrázek 3D). Václav Vopálenský pomocí enzymů hDcp2 a Rai1 prokázal, že část transkriptů všech genů pGKL plasmidů obsahuje klasickou N<sup>7</sup>-methylguanosinovou čepičkovou strukturu typu *cap* 0. V kombinaci s mými výsledky je tedy velice pravděpodobné, že se na N<sup>7</sup>-methylaci guanosinu mRNA čepičky transkriptů pGKL plasmidů *in vivo* podílí právě protein K2ORF3p, ať už samostatně či jako součást proteinového komplexu. Není však známo, zda je protein K2ORF3p pro pGKL plasmidy esenciální právě díky svým enzymatickým aktivitám souvisejícím se syntézou mRNA čepičky. K proteinu K2ORF3p homologní protein D1 viru *Vaccinia* je multifunkční proteinový komplex, který se kromě syntézy mRNA čepičky podílí i na funkcích v iniciaci a terminaci transkripce genů VACV (HARRIS *et al.* 1993; LUO *et al.* 1995; SHATZER *et al.* 2008). Obdobnou multifunkční povahu proteinu K2ORF3p v transkripci pGKL plasmidů nelze spolehlivě vyloučit. Proto jsme se rozhodli cílenou mutagenezí vybraných aminokyselinových zbytků K2ORF3p testovat, zda je přítomnost aktivního centra methyltransferázy K2ORF3p esenciální pro fungování pGKL plasmidů *in vivo*.

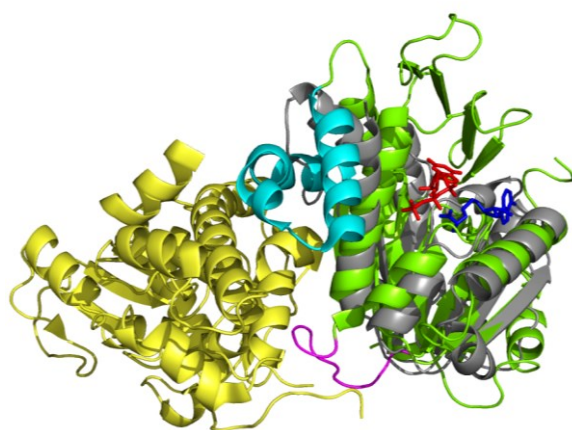
Výsledky ukázaly, že plasmid pRKL2-2 nesoucí bodové mutace v kódující sekvenci methyltransferázy genu *K2ORF3* není schopen v buňkách *K. lactis* zastoupit funkci plasmidu pGKL2, na rozdíl od plasmidu pRKL2-1, jehož sekvence je až na zmíněné bodové mutace

shodná s plasmidem pRKL2-2. Těchto výsledků bylo dosaženo v několika nezávislých klonech získaných po transformaci buněk *K. lactis* IFO1267 a dokonce i v klonech získaných několika nezávislými transformacemi. Ačkoliv rekombinantní plasmidy pRKL2-1 a pRKL2-2 nebyly nikdy sekvenovány v celé délce, ale pouze v oblastech zmíněných zaváděných mutací, téměř jistě jsou za rozdílnou soběstačnost zmíněných plasmidů v buňkách zodpovědné mutace kódující sekvenční metyltransferázové domény K2ORF3p. Jako nejpravděpodobnější vysvětlení nesoběstačnosti plasmidu pRKL2-2 se jeví narušení procesu methylace guanosinové čepičky mRNA pGKL plasmidů, přičemž přítomnost čepičky typu *cap* 0 na transkriptech lineárních plasmidů by tak byla esenciální. Dalším, méně pravděpodobným vysvětlením, by mohlo být zmíněnými mutacemi narušené skládání celé metyltransferázové domény nebo dokonce skládání celého proteinu K2ORF3p, čímž by mohlo dojít k ovlivnění možných jiných esenciálních funkcí proteinu K2ORF3p, nesouvisejících přímo s enzymatickými aktivitami syntézy 5' mRNA čepičky. Protože byly mutovány pouze tři aminokyselinové zbytky proteinu K2ORF3p, považuji tak tuto možnost za vysoce nepravděpodobnou.

Výsledky naznačují, že intaktní metyltransferázová doména K2ORF3p je pro pGKL plasmidy esenciální *in vivo*, pravděpodobně díky své enzymatické funkci. Otázkou ovšem zůstává, proč se nepodařilo metyltransferázovou aktivitu purifikovaného proteinu K2ORF3p detekovat *in vitro*. Možným vysvětlením by mohla být stimulace metyltransferázové aktivity proteinu K2ORF3p jiným proteinem, jako je tomu v případě mRNA *capping* enzymu VACV (HIGMAN *et al.* 1992; MAO a SHUMAN 1994). Proto jsem se bioinformatickými metodami pokusil určit identitu čtyř ORF pGKL plasmidů s dosud neznámou funkcí. Díky poměrně malé molekulové hmotnosti se proteiny K2ORF1p a K2ORF8p jeví jako nejpravděpodobnější kandidáti na případný homolog proteinu D12, což je malá podjednotka mRNA *capping* enzymu VACV (NILES *et al.* 1989). Žádnou podobnost mezi zmíněnými proteiny a D12 se mi najít nepodařilo, zato jsem detekoval podobnost proteinu K2ORF9p se *scaffold* proteinem D13 viru *Vaccinia*.

Na tvorbě komplexu proteinu D1 s proteinem D12 VACV se podílí především aminokyselinové zbytky  $\alpha$ -helixů  $\alpha$ R a  $\alpha$ S proteinu D1 a smyčky mezi  $\alpha$ -helixem  $\alpha$ K a  $\beta$ -listem  $\beta$ 27 proteinu D1 (KYRIELEIS *et al.* 2014). Podle modelu terciární struktury K2ORF3p však protein K2ORF3p tyto strukturní prvky úplně postrádá (Obrázek 71). Je proto vysoce nepravděpodobné, že by metyltransferázová aktivita K2ORF3p mohla být stimulována proteinem, jehož celková struktura by byla podobná proteinu D12 VACV. V souladu s tímto zjištěním nebyl s proteinem K2ORF3p koprecipitován žádný protein pGKL plasmidů s dosud neznámou funkcí, kromě proteinu K2ORF9p, což je již zmíněný předpokládaný *scaffold* protein homologní proteinu D13 VACV. Výsledky reciprokých koimunoprecipitací též naznačují, že K2ORF3p tvoří s nekanonickou RNA polymerázou pGKL plasmidů transkripční komplex

*in vivo*. Interakce proteinu K2ORF3p s RNAP pGKL plasmidů by tak teoreticky mohla být esenciální pro methyltransferázovou aktivitu K2ORF3p, což by vysvětlovalo, proč samostatně purifikovaný protein K2ORF3p tuto aktivitu nevykazuje *in vitro*. Tato situace by byla analogická kvasinkové RNA-guanylyltransferáze Ceg1 *S. cerevisiae*, jejíž enzymatická aktivita je alostericky regulována interakcí s RNA-trifosfatázou Cet1 a s RNA polymerázou II (CHO *et al.* 1998).



**Obrázek 71:** Proložení modelu terciární struktury methyltransferázy mRNA *capping* enzymu K2ORF3p pGKL plasmidů (šedě) s terciární strukturou methyltransferázy mRNA *capping* enzymu D1 viru *Vaccinia* (PDB ID: 4CKB, zeleně) v komplexu s menší podjednotkou mRNA *capping* enzymu D12 viru *Vaccinia* (PDB ID: 4CKB, žlutě) a S-adenosyl-L-homocysteinem (modře). Struktura 7-methylguanosin 5' trifosfátu (červeně) pochází ze struktury mRNA cap (guanin-N7)-methyltransferázy Ecm1 (PDB ID: 1RI1). Detaily zobrazení jsou popsány v legendách Obrázku 3 a 12.  $\alpha$ -helixy  $\alpha$ R a  $\alpha$ S proteinu D1 (tyrkysově) a smyčka mezi  $\alpha$ -helixem  $\alpha$ K a  $\beta$ -listem  $\beta$ 27 proteinu D1 (fialově), podílející se

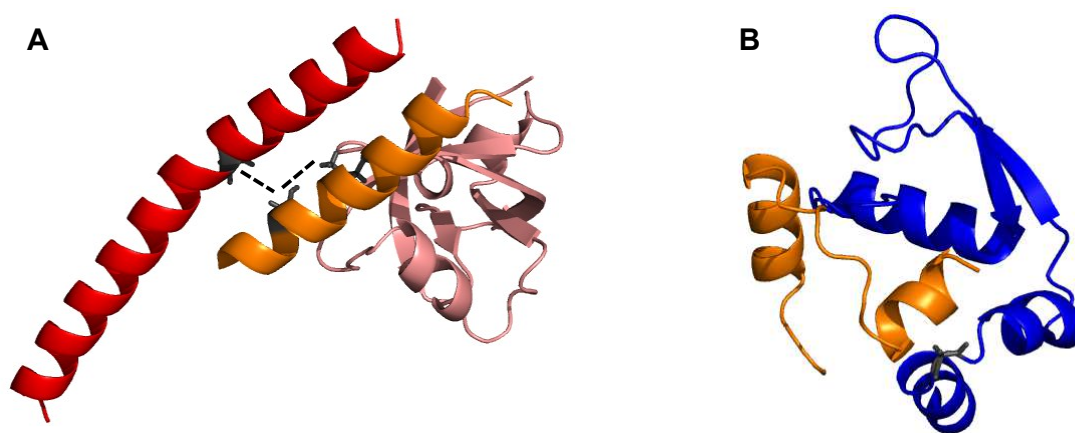
na interakcích s proteinem D12, jsou označeny a pojmenovány podle (KYRIELEIS *et al.* 2014).

### 6.1.2 Složení transkripčního aparátu pGKL plasmidů *in vivo*

Po mnohaletých neúspěšných pokusech o detekci vzájemných interakcí proteinů transkripčního aparátu pGKL plasmidů pomocí různých molekulárně biologických a biochemických přístupů (VOPÁLENSKÝ 1999; SEKYROVÁ 2004; SÝKORA 2013) jsem zavedl a optimalizoval systém umožňující purifikaci proteinů a detekci interakcí proteinů exprimovaných z pGKL plasmidů v buňkách *K. lactis*. Obtížnost exprese AT bohatých plasmidových genů v konvenčních expresních systémech byla obejitá přípravou modifikovaných a dvojité modifikovaných pGKL plasmidů homologní rekombinací *in vivo*. Takto připravené modifikované plasmidy exprimují transkripční komponenty s proteinovými značkami ve svých přirozených fyziologických koncentracích. Použitím koimunoprecipitace, následované hmotnostní spektrometrií a metodou *western blot*, jsem prokázal, že jádro transkripčního komplexu pGKL plasmidů se sestává ze dvou podjednotek RNA polymerázy (K2ORF6p, K2ORF7p) a mRNA *capping* enzymu (K2ORF3p) *in vivo*. Interakce velké podjednotky RNAP a mRNA *capping* enzymu je nezávislá na přítomnosti nukleových kyselin. Následně jsem detekoval, že předpokládaná DExD/H box helikáza (K2ORF4p) je poněkud slaběji asociována jak s velkou podjednotkou plasmidové RNAP, tak s plasmidovým mRNA *capping* enzymem *in vivo*, přičemž interakce K2ORF4p s RNAP je

taktéž nezávislá na přítomnosti nukleových kyselin. Výsledky naznačují níže popsany molekulární model interakce složek transkripčního aparátu pGKL plasmidů.

Zaprvé, proteiny K2ORF6p a K2ORF7p budou s největší pravděpodobností interagovat přímo, neboť je v nich zachováno mnoho univerzálně konzervovaných aminokyselinových zbytků, o nichž je známo, že se účastní intramolekulárních nebo intermolekulárních vazeb ve vícepodjednotkových RNA polymerázách, je v proteinech K2ORF6p a K2ORF7p podle mojí bioinformatické analýzy zachováno. Interagovat bude především konzervovaná oblast  $\beta'$ a16 proteinu K2ORF7p s oblastmi  $\beta'$ a12 a  $\beta'$ a16 proteinu K2ORF6p (Obrázek 72A) a s největší pravděpodobností i konzervovaná oblast  $\beta'$ a20 proteinu K2ORF7p s oblastí  $\beta$ a15 proteinu K2ORF6p (Obrázek 72B).



**Obrázek 72:** Předpokládané přímé interakce podjednotek RNA polymerázy pGKL plasmidů. (A) Interakce oblastí  $\beta'$ a16 (oranžově) proteinu K2ORF7p s oblastmi  $\beta'$ a12 (lososově) a  $\beta'$ a16 (červeně) proteinu K2ORF6p. Konzervované aminokyselinové zbytky L<sup>708</sup>, T<sup>1088</sup> a T<sup>1234</sup>  $\beta'$  podjednotky (číslování podle RNAP *Thermus thermophilus*), podílející se na vzájemných interakcích (LANE a DARST 2010b), jsou u RNAP lineárních plasmidů zachovány. Aminokyselinové zbytky L<sup>815</sup> a T<sup>959</sup> K2ORF6p a T<sup>16</sup> K2ORF7p jsou zobrazeny šedě, přičemž jejich předpokládané vzájemné interakce jsou naznačeny přerušovanou čarou. (B) Interakce oblastí  $\beta'$ a20 (oranžově) proteinu K2ORF7p s oblastí  $\beta$ a15 (modře) proteinu K2ORF6p. Z konzervovaných aminokyselinových zbytků G<sup>1044</sup>  $\beta$  podjednotky a G<sup>1475</sup>  $\beta'$  podjednotky, podílejících se na vzájemných interakcích (LANE a DARST 2010b), je u RNAP lineárních plasmidů zachován aminokyselinový zbytek G<sup>634</sup>  $\beta$  podjednotky u K2ORF6p (šedě), který by teoreticky mohl interagovat s některým aminokyselinovým zbytkem C-koncového  $\alpha$ -helixu K2ORF7p. Strukturní model převzat ze (SÝKORA 2013).

Zadruhé, mRNA *capping* enzym bude nejspíš interagovat přímo s komplexem RNAP, ačkoliv není zřejmé, zda se půjde o interakci s proteinem K2ORF6p a/nebo K2ORF7p. Je známo, že buněčné mRNA *capping* enzymy interagují s C-terminální doménou RNAP II (MCCRACKEN *et al.* 1997). Homologní C-terminální doména však není přítomna ani u RNAP viru *Vaccinia* ani u RNAP lineárních plasmidů. Přesto však heterodimerní mRNA *capping* enzym VACV interaguje přímo s komplexem virové RNAP a předpokládá se, že je přítomen jak při iniciaci transkripce, tak i během elongace transkripce (LUO *et al.* 1991; HAGLER a SHUMAN

1992a). Analogicky může tedy K2ORF3p využívat obdobný způsob interakce s komplexem RNAP pGKL plasmidů. Těsná interakce proteinu K2ORF3p s plasmidovou RNAP také připomíná obdobnou situaci u virů čeledi *Baculoviridae*, kde je virový mRNA *capping* enzym kopurifikován s virovou nekanonickou RNA polymerázou, která má v mnoha ohledech podobnou redukovanou architekturu jako RNAP pGKL plasmidů [pro přehled viz (SÝKORA 2013)], a je chápán jako podjednotka RNAP (GROSS a SHUMAN 1998; GUARINO *et al.* 1998; JIN *et al.* 1998).

Zatřetí, výsledky naznačují méně pevnou vazbu předpokládané helikázy na jádro transkripčního komplexu pGKL plasmidů, než je tomu v případě vzájemných vazeb podjednotek RNAP a mRNA *capping* enzymu. Je známo, že homologní helikázy D6 (součást heterodimerního VETF) a D11 (NPH I) viru *Vaccinia* interagují s virovou RNAP prostřednictvím asociovaného proteinu RAP94, přičemž tyto interakce jsou specifické pro RNAP zabalenou ve virionu, schopnou přepisovat časné geny viru (AHN *et al.* 1994; MOHAMED a NILES 2000; YANG a MOSS 2009). VETF může interagovat s RNAP pouze za přítomnosti RAP94 (YANG a MOSS 2009). Zdá se však, že mimo skupinu poxvirů neexistují homology proteinu RAP94 u ostatních nukleocytoplasmatických virů (BROYLES a KNUTSON 2010), což naznačuje existenci dalšího dosud neznámého mechanismu interakce paralogů proteinů D6 a D11 s virovými RNAP u jiných nukleocytoplasmatických virů. Fylogeneticky příbuzné helikázy typu D6/D11 byly totiž kromě poxvirů a kvasinkových lineárních plasmidů identifikovány i u virů čeledí *Asfarviridae*, *Phycodnaviridae*, *Marseilleviridae* a *Mimiviridae* (YÁÑEZ *et al.* 1993; LU *et al.* 1996; BOYER *et al.* 2009; YUTIN a KOONIN 2012). Dosud neznámý mechanismus interakce paralogů helikáz D6 a D11 s RNAP virů těchto čeledí by se tak mohl analogicky vyskytovat i u pGKL plasmidů. Bohužel kromě modelového poxviru *Vaccinia* nebylo složení transkripčního aparátu žádného jiného nukleocytoplasmatického viru studováno. Důležité však je, že transkripční aparát kvasinkových lineárních plasmidů bude nejspíš úplně soběstačný, protože pomocí hmotnostní spektrometrie nebyly identifikovány žádné hostitelské buněčné proteiny, které by byly prokazatelně specificky spojeny s transkripčním komplexem pGKL plasmidů.

### 6.1.3 Produkce RNA polymerázy pGKL plasmidů v bakteriálním expresním systému

Po mnohaletých poměrně neúspěšných pokusech o produkci a purifikaci RNA polymerázy pGKL plasmidů v bakteriálních expresních systémech využívajících bakterii *E. coli* (VOPÁLENSKÝ 1999; SÝKORA 2013) se ani v této práci nepodařilo dosáhnout lepších výsledků



v produkci rekombinantního proteinu GST-K2ORF6p z připravených duálních expresních vektorů. Experimentální uspořádání duálních vektorů s bicistronními kazetami možná nebylo vhodně zvolené, protože se v lyzátech expresních kmenů po indukci exprese nepodařilo spolehlivě identifikovat protein K2ORF7p-His pomocí metody *western blot* následované imunodetekcí, přičemž s produkcí rekombinantních proteinů K2ORF7p většinou není problém, a tyto proteiny jsou v lyzátech expresních buněk detekovatelné i barvením polyakrylamidových gelů pomocí *Coomassie Blue* (VOPÁLENSKÝ 1999; SÝKORA 2013). Výsledky tak nasvědčují, že v expresních kmenech nedochází k expresi druhého cistronu, a velmi pravděpodobně ani k expresi prvního cistronu, kódovaného duálním vektorem. Nabízí se tak otázka, jestli v expresních buňkách vůbec vzniká dostatek bicistronních transkriptů, či zda není nějaký problém s translací těchto transkriptů.

Nedávno bylo ukázáno, že produkce proteinů pGKL plasmidů v kvasinkových jaderných expresních systémech není možná díky vysokému obsahu AT bází plasmidových genů, který indukuje fragmentaci mRNA těchto genů (KAST *et al.* 2015). Otázkou je, zda při expresi genů pGKL plasmidů s vysokým obsahem AT bází nedochází v bakteriích k nějakému analogickému systémovému problému, který by kompletně znemožňoval produkci příslušných proteinů. V tomto ohledu je zajímavých několik relativně recentních publikací, které na různých bakteriálních modelech demonstrují transkripční umlčení AT bohatých genů *in vivo* vazbou DNA vazebných proteinů, například proteinu H-NS, na AT bohatou DNA (LUCCHINI *et al.* 2006; NAVARRE *et al.* 2006; GORDON *et al.* 2010; SMITS a GROSSMAN 2010). Exprese AT bohaté DNA v buňkách *E. coli* též snižuje *fitness* buněk a způsobuje cytotoxicitu sekvestrací bakteriální RNAP na AT bohatou DNA (RAGHAVAN *et al.* 2012; LAMBERTE *et al.* 2017). Tyto jevy zřejmě souvisí s neúspěchem exprese genu *K2ORF6* v bakteriálních expresních systémech. Zatímco exprese menšího genu *K2ORF7* (399 bp) v *E. coli* je možná a přítomnost vektoru s genem *K2ORF7* v buňkách *E. coli* téměř neovlivňuje růstové vlastnosti těchto buněk, exprese většího genu *K2ORF6* (2925 bp) je téměř nedetekovatelná a růst buněk *E. coli* s vektorem nesoucím gen *K2ORF6* je velmi rychle inhibován při teplotě 37 °C a 20 °C, oproti kontrolním buňkám stejného kmene nesoucím prázdný vektor (VOPÁLENSKÝ 1999).

Společná koexprese genů *K2ORF6* a *K2ORF7* v této práci tak zřejmě zapříčinila, že se nepodařilo produkovat ani protein K2ORF7p, neboť možná transkripční represe genu *K2ORF6* nutně ovlivní druhý cistron duálního expresního vektoru. Pro úspěšnou produkci proteinů RNAP pGKL plasmidů v buňkách *E. coli* by bylo zřejmě třeba použít uměle syntetizované geny s optimalizovaným množstvím GC párů a optimalizovanými kodóny. Podobný experimentální přístup umožnil, do té doby neúspěšnou, produkci imunitní složky několika lineárních cytoplasmatických plasmidů v kvasinkovém jaderném expresním systému (KAST *et al.* 2015).



Díky výše uvedeným dlouhodobým zkušenostem s produkcí K2ORF6p a nedostatečnému rozpočtu pro umělou syntézu upraveného genu *K2ORF6* jsem zanechal veškerých pokusů o produkci kompletní RNAP pGKL plasmidů v bakteriích, neboť osobně považuji úspěch tohoto počínání za nereálný.

#### 6.1.4 Testování aktivit transkripčního aparátu pGKL plasmidů *in vitro* a *in vivo*

Ačkoliv původní cíl mojí diplomové a dizertační práce spočíval v produkci a purifikaci RNAP pGKL plasmidů a její následné charakterizaci *in vitro*, neúspěšná produkce této RNAP v jakémkoliv zvoleném expresním systému způsobila, že jsem se zaměřil na charakterizaci transkripčního aparátu pGKL jako celku pomocí bioinformatických metod a přístupů *in vivo*. Afinitní izolace proteinů transkripčního aparátu pGKL plasmidů označených pomocí GFP z buněk *K. lactis* v detekovatelných množstvích pomocí GFP-Trap®\_A mě však přiměla otestovat některé předpokládané aktivity vybraných proteinů *in vitro*.

Nejprve jsem se pokusil ověřit RNA polymerázovou aktivitu izolovaného transkripčního komplexu pGKL plasmidů *in vitro* pomocí poměrně citlivé detekce inkorporace radioaktivně značených nukleotidů do materiálu nerozpustného v kyselině trichloroctové. První experiment sice naznačoval slabou RNA polymerázovou aktivitu purifikovaného transkripčního komplexu, ale obdobný výsledek se nepodařilo zreprodukovat. Důvodů, proč se nepodařilo detekovat aktivitu předpokládané RNAP, může být celá řada. Jako nejvíc pravděpodobný důvod se jeví zvolení nevhodných podmínek pro enzymatickou aktivitu proteinu (pH, iontová síla, koncentrace a identita dvoumocných kationtů). Tento protokol byl sestaven převážně podle protokolů testujících nespecifickou RNA polymerázovou aktivitu purifikované RNAP viru *Vaccinia in vitro*, která vyžaduje přítomnost  $Mn^{2+}$  iontů, ať už jako jediných dvoumocných kationtů v reakci či v kombinaci s  $Mg^{2+}$  ionty (BAROUDY a MOSS 1980; SPENCER *et al.* 1980). Aktivita RNAP pGKL plasmidů byla testována za současné přítomnosti  $Mn^{2+}$  a  $Mg^{2+}$  iontů, takže nelze vyloučit například inhibiční efekt  $Mn^{2+}$  iontů na aktivitu plasmidové RNAP. Iontová síla transkripční reakce se většinou nastavuje pomocí 100 – 150 mM KCl. V použitém protokolu byla použita nižší koncentrace KCl (50 mM), neboť elongační komplexy bakteriální RNAP postrádající strukturní prvek kormidlo (v originále *rudder*) jsou méně stabilní ve zvýšených koncentracích KCl (KUZNEDELOV *et al.* 2002), přičemž tento strukturní prvek podle mojí bioinformatické analýzy RNA polymeráza pGKL plasmidů postrádá (SÝKORA 2013). Aktivita RNAP pGKL plasmidů *in vitro* by tak teoreticky mohla vyžadovat ještě nižší koncentraci KCl než použitou.

Neúspěch při detekci aktivity předpokládané RNAP však může mít i jiný důvod. Pro reakci byly použity NTP ze soupravy TranscriptAid T7 High Yield Transcription Kit (Thermo Fisher Scientific), jež měla cca 2 roky po doporučeném datu spotřeby. Přibližně rok po mých pokusech *in vitro* vyšlo najevo, že zmíněná souprava nefunguje, přičemž výměna NTP soupravy za NTP z nové soupravy způsobila její opětovnou funkčnost. Protože moje pokusy postrádaly jakoukoliv jinou RNAP jako pozitivní kontrolu reakce, nelze vyloučit, že se nepodařilo detekovat RNA polymerázovou aktivitu transkripčního komplexu pGKL plasmidů díky degradovaným substrátům.

Poté jsem se pokusil ověřit DNA vazebnou aktivitu izolovaného transkripčního komplexu a izolované předpokládané helikázy pGKL plasmidů *in vitro* pomocí poměrně citlivé detekce imobilizace radioaktivně značené DNA imunoprecipitovanými proteiny, která byla adaptována podle (KIM 2004). Pro experimenty byly vybrány dva úseky dvouvlákové DNA pGKL plasmidů obsahující i neobsahující promotory. Ani v jednom z experimentů se nepodařilo detekovat vazbu DNA vybranými proteiny transkripčního aparátu pGKL plasmidů. Důvodů, proč se nepodařilo detekovat DNA vazebnou aktivitu vybraných proteinů, může být celá řada.

Jako pravděpodobný důvod se jeví zvolení nevhodných podmínek pro reakci *in vitro* (pH, iontová síla, koncentrace a identita dvoumocných kationtů). Reakce probíhala za neutrálního pH a nízké iontové síly. Interakce proteinů VETF a NPH I viru *Vaccinia* s DNA *in vitro* jsou poměrně silné; dochází k nim za neutrálních pH, nízké iontové síly a nevyžadují přítomnost dvoumocných kationtů, ačkoliv přítomnost  $Mg^{2+}$  může v případě faktoru VETF jeho vazbu na DNA *in vitro* zesilovat (YUEN *et al.* 1987; BROYLES *et al.* 1988; HAGLER a SHUMAN 1992c; DENG a SHUMAN 1998; CHRISTEN *et al.* 1998). K sekvenčně nespecifické interakci purifikované RNAP VACV s DNA *in vitro* také dochází za neutrálního pH a nízké iontové síly bez přítomnosti dvoumocných kationtů, ale RNAP je takto schopná interagovat pouze s ssDNA a ne s dvouvláknovou DNA (SPENCER *et al.* 1980). K sekvenčně specifické interakci purifikované RNAP VACV s dvouvláknovou DNA *in vitro* je potřebná přítomnost transkripčních faktorů (HAGLER a SHUMAN 1992c; LI a BROYLES 1993b). Analogicky tak možná purifikovaná RNAP pGKL plasmidů není sama o sobě schopna vázat dvouvláknovou DNA *in vitro*.

Na základě mojích sekvenčních analýz promotorů pGKL plasmidů, naznačující příbuznost s promotory časných genů poxvirů, jsem ale předpokládal, že purifikovaná helikáza K2ORF4p bude specificky interagovat s dvouvláknovou DNA pGKL plasmidu obsahující promotor *in vitro*, jako je tomu analogicky u faktoru VETF VACV (YUEN *et al.* 1987; BROYLES *et al.* 1991). Pozdější výsledky dokládající specifickou interakci K2ORF4p s DNA pGKL plasmidů *in vivo* tedy nasvědčují, že DNA vazebné pokusy *in vitro* byly neúspěšné, alespoň v případě purifikované helikázy, vlivem technického provedení a ne absencí testované aktivity proteinu. Stejně jako

v případě testování RNA polymerázové aktivity purifikované RNAP pGKL plasmidů se tedy projevil negativní efekt absence pozitivní kontroly DNA vazebné reakce *in vitro* na vyvození hlubších závěrů ze získaných výsledků.

Po neúspěšných pokusech o detekci DNA vazebné aktivity předpokládané helikázy *in vitro* jsem se rozhodl testovat její předpokládané spojení s DNA *in vivo* pomocí chromatinové imunoprecipitace, neboť Josef Novák tou dobou tuto metodu v naší laboratoři zaváděl. Po prvotním neúspěchu, způsobeném pravděpodobně použitím příliš stringentních promývacích pufrů, se nám podařilo detekovat interakci K2ORF4p s DNA. Asociace helikázy s DNA byla specifická pro DNA pGKL plasmidů, což je v souladu s předpokládanou rolí K2ORF4p v přímé interakci s promotorovým elementem UCS pGKL plasmidů *in vivo*. Pomocí chromatinové imunoprecipitace se však nepodařilo blíže identifikovat vazebné místo proteinu K2ORF4p, neboť ve vzorku chromatinového imunoprecipitátu byly shodně nabohaceny jak úseky DNA pGKL plasmidů obsahující, tak neobsahující promotorové elementy UCS. Vzhledem ke kompaktnosti genomu pGKL plasmidů a malé vzdálenosti mezi jednotlivými UCS motivy je však zřejmě metoda chromatinové imunoprecipitace, využívající sekvenčně nespecifickou fragmentaci DNA pomocí sonikace, pro mapování vazebných míst K2ORF4p ne zcela vhodná.

Pro mapování vazebných míst K2ORF4p *in vivo* by zřejmě šla použít relativně nová metoda ChIP-exo, která kombinuje chromatinovou imunoprecipitaci se štěpením DNA pomocí exonukleázy bakteriofága  $\lambda$ , což zvyšuje rozlišovací schopnost mapování vazebných míst ze stovek bp na jednotky bp (RHEE a PUGH 2011). Analogicky jsem se pokusil vyvinout vlastní metodu ChIP-endo, která by kombinovala klasickou chromatinovou imunoprecipitaci využívající sonikaci s dalším stupněm fragmentace DNA pomocí restričních endonukleáz typu II. Štěpení crosslinkované DNA navázané na imunoprecipitovaný protein K2ORF4p pomocí vybraných restričních endonukleáz se mi však během dvou pokusů nepodařilo optimalizovat (data neuvedena). Podrobnější analýza literatury však ukázala, že obdobná metoda byla úspěšně použita na modelu buněk savčích tkáňových linií, přičemž kompletní restriční štěpení DNA izolovaných jader buněk vyžadovala stovky jednotek restričních endonukleáz a několikahodinové inkubace (SCHUCH *et al.* 2012). Další optimalizace metody ChIP-endo by tedy mohla vést k mapování vazebných míst K2ORF4p *in vivo*, ačkoliv pokud je například DNA pGKL plasmidů v buňkách obalena nějakým proteinem, restriční štěpení crosslinkované plasmidové DNA by nemuselo být technicky možné vlivem nedostupnosti povrchu sekvencí DNA pro interakci s restričním enzymem.

### 6.1.5 Architektura RNA polymerázy pGKL plasmidů

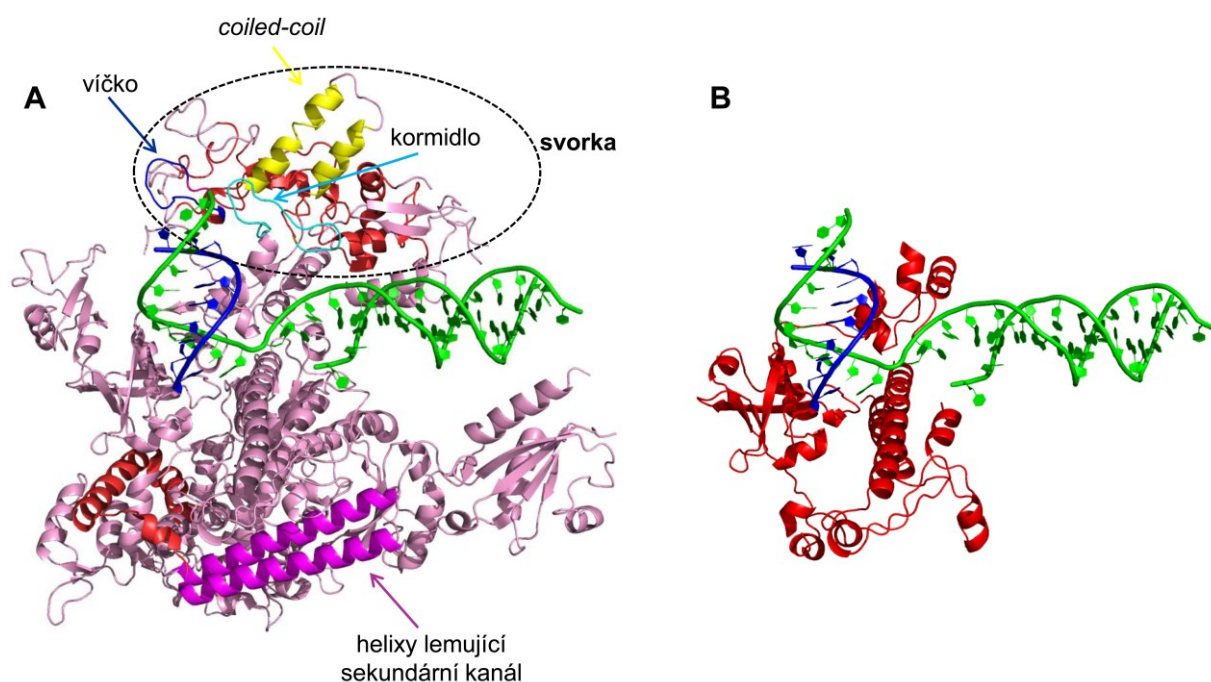
Bioinformatická analýza RNA polymeráz kvasinkových cytoplasmatických plasmidů se ukázala jako poměrně náročná kvůli jedinečné redukované architektuře enzymu a značné sekvenční odlišnosti od jiných vícepodjednotkových RNAP. Z vytvořeného strukturního modelu RNAP pGKL plasmidů je zřejmé, že plasmidové RNAP se významně liší od kanonických RNAP v několika aspektech.

Zaprvé, plasmidové RNAP postrádají téměř celou svorku. Zachována je pouze bazální část svorky, tvořená konzervovanými oblastmi  $\beta$ a15,  $\beta$ a16 a  $\beta$ 'a20. Svorka je mobilní prvek RNAP a její uzavření je důležité pro vysokou stabilitu a procesivitu enzymu. Konformace svorky je u buněčných RNAP regulována interakcí univerzálně konzervovaných elongačních faktorů NusG a Spt4/5 s *coiled-coil* motivem svorky (HIRTREITER *et al.* 2010), který v plasmidové RNAP pravděpodobně chybí (Obrázek 73). Proto je velmi nepravděpodobné, že by transkripční aparát pGKL plasmidů mohl využívat hostitelské proteiny Spt4/5 ke zvýšení procesivity.

Zadruhé, plasmidové RNAP pravděpodobně postrádají víčko a kormidlo (Obrázek 73). Víčko RNAP působí jako klín usnadňující dislokaci nascentní RNA z DNA:RNA hybridní molekuly, čímž se podílí na udržování konstantní délky DNA:RNA hybridní molekuly mezi 7 až 10 páry bází (VASSYLYEV *et al.* 2007). Absence víčka u bakteriální a archeální RNAP negativně ovlivňuje transkripci ssDNA, přičemž transkripce dvouvláknových DNA templátů těmito modifikovanými RNAP není ovlivněna *in vitro* (TOULOKHONOV a LANDICK 2006; NAJI *et al.* 2008). Analogicky by tak dislokace nascentní RNA z DNA:RNA hybridní molekuly *in vivo* neměla být ovlivněna nepřítomností víčka v plasmidové RNAP, přepisující dvouvláknovou DNA lineárních plasmidů. Byla také navržena hypotéza o vlivu víčka RNAP na vnitřní terminaci transkripce využívající RNA vlásenky u bakterií (VASSYLYEV *et al.* 2007). Jelikož však je bakteriální RNAP bez víčka schopná vnitřní terminace transkripce *in vitro* (TOULOKHONOV a LANDICK 2006), předpokládám, že tento strukturní prvek RNAP není esenciální pro předpokládanou vnitřní terminaci transkripce lineárních plasmidů *in vivo*.

Kormidlo RNAP interaguje s horním okrajem DNA:RNA hybridní molekuly (VASSYLYEV *et al.* 2007). Experimenty používající bakteriální RNAP s chybějícím kormidlem odhalily defekty při iniciaci transkripce a menší stabilitu elongačních komplexů RNAP *in vitro* (KUZNEDELOV *et al.* 2002). Předpokládaná menší stabilita elongačních komplexů RNAP lineárních plasmidů oproti elongačním komplexům kanonických RNAP by mohla vysvětlit možnou terminaci transkripce lineárních plasmidů indukovanou poměrně slabými RNA vlásenkami. Kormidlo se nachází mezi konzervovanými oblastmi  $\beta$ 'a10 a  $\beta$ 'a11 a mezi jednotlivými třídami RNAP není sekvenčně konzervované (LANE a DARST 2010b). Nedá se tak

vyložit, že insert mezi oblastmi  $\beta$ a16 a  $\beta'$ a11 RNAP pGKL plasmidů nebude alespoň částečně plnit roli kormidla kanonických RNAP. Důležitost insertu mezi oblastmi  $\beta$ a16 a  $\beta'$ a11 RNAP pGKL plasmidů pro funkci RNAP *in vivo* dokládá pokus o rozdělení fúzovaných oblastí  $\beta$  a  $\beta'$  podjednotky RNAP v proteinu K2ORF6p, kde plasmid pRKL2-19 nebyl schopen v buňkách zastoupit funkci plasmidu pGKL2. To značí, že insert mezi oblastmi  $\beta$ a16 a  $\beta'$ a11 není pouhým spojením  $\beta$  a  $\beta'$  podjednotky plasmidové RNAP, které by spolu měly interagovat i v případě rozdělení K2ORF6p na dva polypeptidy, ale má pravděpodobně i další funkci. Produkce proteinů K2ORF6 $\beta$ p a K2ORF6 $\beta'$ p v kmeni *K. lactis* IFO1267\_pRKL2-19 stejně tak jako předpokládaná interakce proteinů K2ORF6 $\beta$ p a K2ORF6 $\beta'$ p však testována nebyla.



**Obrázek 73:** Redukovaná architektura RNA polymerázy pGKL plasmidů. (A) Krystalová struktura Rpb1 podjednotky (ortholog  $\beta'$  podjednotky, růžově) elongačního komplexu RNAP II *S. cerevisiae* (PDB ID: 2NVQ) v komplexu s DNA (zeleně) a RNA (modře). Šípky označují důležité strukturní prvky  $\beta'$  podjednotky, které se nachází u všech kanonických vícepodjednotkových RNAP, a které RNAP pGKL plasmidů patrně postrádá. Jedná se o víčko (modře), *coiled-coil* (žlutě), kormidlo (tyrkysově) a helixy lemující sekundární kanál (fialově). Mimo popsanych strukturních prvků jsou zobrazeny i další konzervované oblasti  $\beta'$  podjednotky, jež nejsou zachovány u RNAP pGKL plasmidů (tmavě červeně). (B) Model terciární struktury  $\beta'$  podjednotky RNA polymerázy pGKL plasmidů (červeně). Model ukazuje aminokyselinové zbytky 754–882 a 894–974 K2ORF6p, a 1–52 a 103–132 K2ORF7p, DNA (zeleně) a RNA (modře). Nukleové kyseliny modelu jsou založeny na strukturách nukleových kyselin elongačního komplexu RNAP II *S. cerevisiae* (PDB ID: 2NVQ). Model byl se strukturou RNAP II porovnán pomocí DalíLite v. 3 a vizualizován pomocí PyMOL 1.3.

Zatřetí, plasmidové RNA polymeráze chybí helixy lemující sekundární kanál (Obrázek 73). Tyto dva  $\alpha$ -helixy jsou vazebnými místy pro některé transkripční faktory vícepodjednotkových RNAP, jako je například transkripční elongační faktor TFIIIS



(KETTENBERGER *et al.* 2004). Proto je vysoce nepravděpodobné, že by transkripční aparát lineárních plasmidů mohl používat buněčný TFIIS hostitelských kvasinkových buněk ke zvýšení své opravné štěpící aktivity. V souladu s bioinformatickou analýzou RNAP pGKL plasmidů nebyly pomocí hmotnostní spektrometrie nalezeny žádné iniciační a elongační transkripční faktory hostitelských buněk, které by specificky interagovaly s transkripčním komplexem pGKL plasmidů *in vivo*.

### 6.1.6 Evoluční původ transkripčního aparátu pGKL plasmidů

Virový původ genů lineárních kvasinkových plasmidů kódujících mRNA *capping* enzym a předpokládanou helikázu byl navržen již dříve (JESKE *et al.* 2007). Virový původ plasmidových genů kódujících RNAP však nebyl očekáván, neboť předchozí hypotézy naznačovaly, že tyto geny pocházejí z genů RNAP hostitelských kvasinkových buněk (JESKE *et al.* 2007; RUPRICH-ROBERT a THURIAUX 2010), nebo že jsou tyto geny dávnými zástupci diverzifikace vícepodjednotkových RNAP díky jednoduché architektuře těchto RNAP (IYER a ARAVIND 2012). Pro podporu výše uvedených hypotéz však nebyla provedena žádná fylogenetická analýza. Moje fylogenetická analýza naznačuje, že geny RNAP lineárních plasmidů kvasinek mají shodný původ s geny RNAP nukleocytoplasmatických virů.

Nukleocytoplasmatické viry všech čeledí kódují virovou RNA polymerázu sloužící k transkripci virových genů v cytoplasmatické části životního cyklu viru, přičemž tato RNAP obsahuje minimálně tři podjednotky orthologní k  $\beta$ ,  $\beta'$  a Rpb5 podjednotce (IYER *et al.* 2006). Výjimku tvoří většina virů čeledi *Phycodnaviridae*, které ztratily geny kódující RNAP a transkripce jejich genů probíhá pouze v jaderné části jejich životního cyklu činností hostitelského transkripčního aparátu (IYER *et al.* 2006). Celkové množství podjednotek virových RNAP a jejich původ je však u jednotlivých virových čeledí značně variabilní [pro přehled viz (YUTIN a KOONIN 2012; SÝKORA 2013; MIRZAKHANYAN a GERSHON 2017)]. Předpokládá se, že původní orthology  $\beta$  a  $\beta'$  podjednotky virové RNAP společného předka všech nukleocytoplasmatických virů měly monofyletický původ v eukaryotické RNA polymeráze I (YUTIN a KOONIN 2012), přičemž  $\beta$  a  $\beta'$  podjednotky RNAP poxvirů a cytoplasmatických plasmidů se shodně větví právě poblíž eukaryotické RNAP I. Poxvirová RNAP též neobsahuje zřetelné orthology  $\alpha$  podjednotek (KNUTSON a BROYLES 2008), stejně jako plasmidová RNAP, a je celkově jednodušší než eukaryotické RNA polymerázy. Proto by byl vznik RNAP lineárních plasmidů redukcí virové RNAP mnohem snazší než redukcí kvasinkové RNAP. Moje fylogenetická analýza tak ukazuje, že všechny geny transkripčního aparátu



lineárních kvasinkových plasmidů mají shodný původ s geny nukleocytoplasmatických virů, blízký genům virů čeledi *Poxviridae*.

Tyto výsledky jsou ve shodě s analýzou mRNA *capping* enzymů lineárních plasmidů, kde fylogenetická analýza, založená na konzervovaných motivech guanylyltransferázy, také naznačila poxvirový původ těchto genů (JESKE *et al.* 2007). Moje fylogenetická analýza mRNA *capping* enzymů lineárních plasmidů, založená na sekvencích aminokyselinových zbytků methyltransferáz, jež mají být u nukleocytoplasmatických virů monofyletické (YUTIN a KOONIN 2012), však poukázala na původ blízký virům čeledi *Marseilleviridae* (data nezobrazena), obdobně jako práce SATWIKA *et al.* (2012). Tripartitní mRNA *capping* enzym VACV se liší od dalšího strukturně charakterizovaného tripartitního enzymu mRNA *capping* enzymu viru čeledi *Mimiviridae* absencí motivu B trifosfatázy (KYRIELEIS *et al.* 2014). Navzdory předchozím sekvenčním analýzám, poukazujícím na možnou přítomnost motivu B trifosfatázy u mRNA *capping* enzymu VACV a lineárních kvasinkových plasmidů (LARSEN *et al.* 1998; PEI *et al.* 1999), se mi pomocí strukturního modelování proteinu K2ORF3p nepodařilo motiv B v mRNA *capping* enzymu lineárních plasmidů identifikovat. Předpokládám tak, že v mRNA *capping* enzymech lineárních plasmidů není motiv B trifosfatázy stejně jako u proteinu D1 VACV přítomen (KYRIELEIS *et al.* 2014). V sekvencích mRNA *capping* enzymů virů čeledi *Marseilleviridae* se mi podařilo motiv B trifosfatázy identifikovat jednoduchým BLASTP<sup>®</sup> hledáním (data nezobrazena). Předpokládám proto, že geny mRNA *capping* enzymů lineárních plasmidů mají, stejně jako geny RNAP lineárních plasmidů, nejbližší právě ke genům virů čeledi *Poxviridae*.

Moje sekvenční analýza protomotorů pGKL plasmidů naznačuje, že nejen plasmidová RNAP, ale také plasmidové promotory jsou fylogeneticky příbuzné promotorům nukleocytoplasmatických virů. Zaznamenal jsem sekvenční podobnost mezi motivem UCE časných genů VACV a rozšířeným motivem UCS pGKL plasmidů, stejně jako podobné umístění těchto motivů vzhledem k místu transkripčního startu. Motiv s podobnou sekvencí jako UCE motiv virů *Poxviridae* byl také identifikován u promotorů časných genů virů čeledi *Mimiviridae*, ačkoliv umístění tohoto motivu vzhledem k místu transkripčního startu se poněkud liší (SUHRE *et al.* 2005). Promotory časných genů ostatních nukleocytoplasmatických virů jsou také poměrně krátké a AT bohaté, ačkoliv nejsou tak podrobně sekvenčně a funkčně charakterizované, jako právě promotory poxvirů [pro přehled viz (OLIVEIRA *et al.* 2017)].

Navzdory značné sekvenční podobnosti UCE a UCS promotorových motivů (Obrázek 39) nefungoval časný promotor genu *D12L* VACV jako promotor po umístění do pGKL plasmidu. To může mít vysvětlení v částečných sekvenčních rozdílech promotorů časných genů poxvirů a lineárních plasmidů. Zatímco konsenzus motiv UCE vykazuje na jednotlivých pozicích jistou variabilitu nukleotidů (YANG *et al.* 2010; YANG *et al.* 2011a), konsenzus motiv UCS

vykazuje invariantní ATNTGA hexanukleotidovou sekvenci, která je univerzální napříč všemi kvasinkovými cytoplasmatickými plasmidy (STARK *et al.* 1990; HISHINUMA a HIRAI 1991; FUKUDA *et al.* 1997; KLASSEN *et al.* 2002; KLASSEN a MEINHARDT 2003; JESKE a MEINHARDT 2006). Přes značnou sekvenční podobnost promotorů časných genů poxvirů a lineárních plasmidů však zřejmě neexistuje žádný časný promotor VACV, který by na pozici odpovídající UCS obsahoval ATNTGA hexanukleotid (YANG *et al.* 2010). Proto zřejmě promotor genu *D12L* VACV, obsahující v rámci UCE sekvence hexanukleotid TAATGA, nefungoval jako promotor po umístění do pGKL plasmidu.

Požadavek na striktní hexanukleotidový konsenzus motiv promotoru místo na méně striktní delší AT bohatý motiv se mohl u lineárních plasmidů vyvinout v rámci minimalizace všudypřítomné iniciace transkripce, ke které by v případě rozeznávání UCE elementu VACV zřejmě docházelo. Pravděpodobnost náhodného výskytu AT bohatého UCE motivu je totiž u pGKL plasmidů (~74 % obsah AT bází) vyšší než u VACV (~67 % obsah AT bází). I tak je četnost výskytu UCE motivu u VACV poměrně vysoká. Bylo ukázáno, že pouze UCE elementy následované desetinukleotidovou AT bohatou sekvencí (medián 80 % AT) jsou schopny indukovat iniciaci transkripce *in vivo* (YANG *et al.* 2011a). U promotorů pGKL plasmidů má deset nukleotidů následujících rozšířený UCS motiv medián obsahu AT bází 75 % (data nezobrazena), což přibližně odpovídá celkovému obsahu AT bází pGKL plasmidů (~74 %). Konsenzus motiv UCE VACV v kombinaci s následující AT bohatou oblastí by tak v rámci genomu lineárních plasmidů nepřispíval k sekvenční specifičtější promotorů v takové míře, jaké je dosaženo v rámci genomu VACV. Otázkou zůstává, zda by některý promotor časného genu VACV s mutacemi UCE motivu, upravujícími sekvenci na pozici odpovídající motivu UCS na ATNTGA, dokázal v pGKL plasmidech indukovat iniciaci transkripce *in vivo* na stejném iniciačním nukleotidu jako u VACV, což by alespoň částečně poukázalo na obdobnou funkci pravděpodobně fylogeneticky příbuzných promotorů.

### 6.1.7 Iniciace transkripce a složení 5' konců mRNA pGKL plasmidů

Předchozí 5' RACE-PCR experimenty mého školitele odhalily krátké poly(A) leadery na 5' koncích mRNA většiny genů pGKL plasmidů. Tyto 5' poly(A) leadery mRNA byly mezi jednotlivými transkripty heterogenní v délce (1 – 21 adenosinů) a nekomplementární k templátové DNA. Netemplátové 5' poly(A) leadery jsou charakteristickým znakem transkriptů středních a pozdních genů viru *Vaccinia* a vznikají v důsledku klouzání RNA polymerázy na iniciátorovém regionu promotoru (BERTHOLET *et al.* 1987; SCHWER *et al.* 1987; SCHWER a STUNNENBERG 1988; DAVISON a MOSS 1989b). Pomocí bioinformatické analýzy se mi

podarilo identifikovat předpokládaný INR element TAAAN v promotorech všech genů pGKL plasmidů, většina jejichž transkriptů obsahovala netemplátové adenosiny na 5' konci. Použitím nově vytvořených kmenů *K. lactis* a 5' RACE-PCR analýzy se mi podařilo potvrdit, že přítomnost 5' poly(A) leaderu transkriptů pGKL plasmidů je závislá na sekvenci předpokládaného INR elementu *in vivo*, přičemž netemplátové nukleotidy jsou do vlákna RNA patrně zařazovány při iniciaci transkripce stejným mechanismem, jaký byl navržen pro postreplikativní transkripty VACV, neboť při pokusech s mutacemi INR elementu pozdního promotoru VACV byl detekován obdobný vliv na délku 5' poly(A) leaderu mRNA *in vivo* (DAVISON a MOSS 1989b). V souladu s mechanismem vzniku 5' poly(A) leaderu mRNA pGKL plasmidů klouzáním RNAP na INR elementu jsou i detekované rozdíly v délce netemplátové části 5' poly(A) leaderu transkriptu *in vivo* v závislosti na počtu následných adenosinů INR elementu, přičemž tyto rozdíly byly mezi některými třídami promotorů statisticky signifikantní. Tyto výsledky tak naznačují, že vzrůstající počet adenosinů INR elementu podporuje opakované klouzávání RNA polymerázy na promotoru. Iniciátorový element promotoru pGKL plasmidů je patrně samostatný funkční prvek, protože míra netemplátové 5' polyadenylace mRNA nebyla ovlivněna kódující sekvencí genu *in vivo*.

Nezávisle na výsledcích mého školitele se mi také podařilo potvrdit, že transkripty pGKL plasmidů vznikající *in vivo* obsahují 5' blokuující strukturu, přičemž tato blokuující struktura je přítomna na transkriptech genů, jejichž mRNA většinou obsahuje i neobsahuje netemplátový 5' poly(A) leader. Vzhledem k dalším výsledkům mého školitele je 5' blokuující struktura mRNA s největší pravděpodobností čepička typu *cap* 0. K syntéze 5' čepičky na transkriptech časných genů VACV *in vitro* dochází kotranskripčně, když je nascentní RNA dlouhá alespoň 31 nukleotidů (HAGLER a SHUMAN 1992a). Analogicky by tak k syntéze 5' čepičky na transkriptech lineárních plasmidů mohlo docházet kotranskripčně *in vivo*, neboť mRNA *capping* enzym K2ORF3p tvoří komplex s RNAP pGKL plasmidů *in vivo*. Vzhledem k celkové redukci RNAP pGKL plasmidů a také díky tomu, že RNAP pGKL plasmidů postrádá víčko, které tvoří část povrchu proximální části RNA odchodného kanálu RNAP (LANE a DARST 2010b), by však teoreticky mohl být povrch 5' konce nascentní RNA pGKL plasmidů dostupný jako substrát pro K2ORF3p i u transkriptů kratších než 31 nukleotidů.

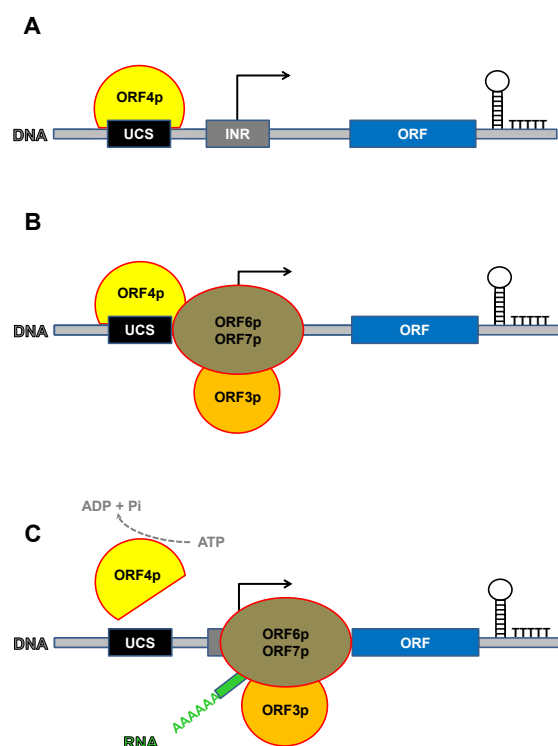
Výsledky z kmene *K. lactis* divokého typu, stejně jako výsledky z nově připravených kmenů s mutovaným INR elementem promotoru, překvapivě ukázaly na signifikantní negativní korelaci mezi počtem netemplátových adenosinů 5' poly(A) leaderu a přítomností 5' mRNA čepičky na transkriptech pGKL plasmidů vznikajících *in vivo*. Obdobný jev se mi později podařilo detekovat i u transkriptů VACV vznikajících *in vivo*, tudíž tento jev není závislý na kódující sekvenci genu ani na prostředí hostitelských buněk. Otázkou je, zda je detekovaný

jev biologicky relevantní. Teoreticky by mohla být schopnost použité reverzní transkriptázy překonat 5'–5' vazbu mezi mRNA čepičkou a prvním nukleotidem mRNA zhoršená u transkriptů s delšími 5' poly(A) sekvencemi. Osobně se ale domnívám, že existuje přímý negativní vztah mezi 5' polyadenylací mRNA a procesem přidání 5' mRNA čepičky u transkriptů lineárních plasmidů a poxvirů *in vivo*, čemuž nasvědčuje i signifikantní rozdíl v délce 5' poly(A) leaderu všech transkriptů pozdního genu *C3L* VACV oproti transkriptům stejného genu obsahujícím 5' mRNA čepičku, izolovaným pomocí metody *oligocapping*. Přepodkládaný negativní vztah by mohl spočívat v tom, že 5' polyadenylované transkripty by fungovaly jako horší substrát pro mRNA *capping* enzym lineárních plasmidů a poxvirů. Tuto hypotézu jsme chtěli otestovat *in vitro* s použitím komerčně dostupného mRNA *capping* enzymu VACV. RNA polyadenylovanou na 5' konci jsme připravili pomocí transkripčního systému *in vitro* s využitím RNA polymeráz SP6 a T7 a modifikovaných promotorů, které umožňují i těmto RNAP klouzat po templátu a přidávat do RNA netemplátové nukleotidy (MILLIGAN a UHLENBECK 1989; EUN 1996). V pilotním experimentu se však nepodařilo připravit RNA *in vitro*, jejíž 5' poly(A) sekvence by byly dostatečně heterogenní v délce, čímž by simulovaly 5' konce mRNA lineárních plasmidů a poxvirů (data nezobrazena). Zřejmě by bylo teoreticky možné připravit RNA dostatečně heterogenní v délce 5' poly(A) sekvencí těmito transkripčními systémy *in vitro* s využitím dvou a více templátů s různým počtem adenosinových zbytků v promotorové oblasti netemplátového vlákna DNA. Tyto experimenty však již provedeny nebyly. Přepodkládaný negativní vztah by však mohl také spočívat v tom, že 5' poly(A) sekvence transkriptů lineárních plasmidů a poxvirů by mohly být v buňkách rychle sekvestrovány vazebnými proteiny, čímž by byl povrch 5' konců těchto transkriptů hůř dostupný jako substrát pro mRNA *capping* enzym.

Vzhledem k předpokládané fylogenetické příbuznosti UCS elementu promotoru lineárních plasmidů s promotorovým elementem UCE časných genů poxvirů a zřejmě i fylogenetické příbuznosti INR elementu promotoru lineárních plasmidů s INR elementem poxvirových genů tak většina promotorů pGKL plasmidů připomíná promotory časných genů poxvirů s INR elementem. Původně byly identifikovány pouze dva promotory časných genů VACV obsahující INR element (AHN *et al.* 1990b; INK a PICKUP 1990). V dnešní době je známo již 18 promotorů časných genů VACV s INR elementem (YANG *et al.* 2011a), což znamená, že přibližně desetina všech známých promotorů genů VACV má obdobné sekvencní složení jako většina promotorů pGKL plasmidů. Rozhodl jsem se tedy pomocí metody RACE-PCR charakterizovat 5' konce mRNA dvou vybraných časných genů VACV, jejichž promotor obsahuje kanonický TAAAT iniciátorový region, a porovnat získané sekvence s 5' konci mRNA genů pGKL plasmidů, jejichž promotor obsahuje TAAAB iniciátorový region. Výsledky ukázaly, že složení 5' konců

porovnávaných mRNA je dost podobné, jak s ohledem na celkovou délku 5' poly(A) leaderu, tak s ohledem na frekvenci výskytu 5' mRNA čepičkové struktury (Tabulka 42). Frekvence výskytu 5' mRNA čepičkové struktury u transkriptů vznikajících z promotorů bez INR je u genů VACV a pGKL plasmidů taktéž podobná (Tabulka 42). Sekvenčně podobné promotory lineárních plasmidů a poxvirů tedy dávají vznik transkriptům s obdobnými 5' konci *in vivo*. Díky tomu, spolu s výsledky dokládajícími interakci předpokládané helikázy pGKL plasmidů s plasmidovou DNA a plasmidovým transkripčním komplexem *in vivo*, jsem navrhl první model klíčových aspektů iniciace a terminace transkripce kvasinkových cytoplasmatických plasmidů (Obrázek 74).

Tabulka 42: Četnost 5' mRNA čepičkových struktur a celková délka 5'poly(A) leaderu transkriptů pGKL plasmidů a viru <i>Vaccinia</i>			
Geny (INR)	Počet sekvencí	Frekvence výskytu 5' čepičky	Medián (průměr) celkové délky 5' poly(A) leaderu
pGKL geny (TAAAB)	117	32 %	7 (7)
časné geny VACV (TAAAT)	32	50 %	5 (5)
pGKL geny bez INR	31	87 %	–
časné geny VACV bez INR	59	86 %	–



**Obrázek 74:** Model iniciace a terminace transkripce lineárních plasmidů kvasinek. **(A)** Předpokládaná helikáza K2ORF4p (ORF4p, žlutě) se váže na plasmidovou DNA, pravděpodobně na konzervovanou sekvenci proti směru transkripce (UCS, černě), která je příbuzná promotorovému elementu UCE časných genů poxvirů. **(B)** K2ORF4p rekrutuje komplex RNAP (ORF6p/ORF7p, hnědě) do místa iniciace transkripce, které obvykle obsahuje iniciátorový region (INR, šedě) zodpovědný za klouzánií RNA polymerázy spojené s 5' polyadenylací mRNA. **(C)** Hydrolýza ATP uvolní K2ORF4p z preiniciačního komplexu a umožní RNAP únik z iniciačního místa promotoru a produkci mRNA (RNA, zeleně) obsahující 5' poly(A) leader. Na tuto RNA může být následně přidána 5' čepička mRNA *capping* enzymem K2ORF3p (ORF3p, oranžově). Terminace transkripce s největší pravděpodobností probíhá mechanismem nezávislým na přídatných faktorech, který zahrnuje vnitřní terminátory, sestávající se ze struktur RNA vlásenek a 3' oligo(U) konce.

Předložený model iniciace transkripce předpokládá interakci proteinu K2ORF4p pGKL plasmidů s UCS elementem promotoru *in vivo*, neboť heterodimerní VETF faktor VACV (obsahuje protein D6, což je předpokládaný homolog K2ORF4p) interaguje s UCE elementem promotoru *in vitro* (YUEN *et al.* 1987; BROYLES *et al.* 1988). Experimenty využívající interferenci DNA–proteinové vazby pomocí methylace DNA ukazují, že VETF interaguje s konzervovaným guanosinovým zbytkem na pozici 7 UCE elementu *in vitro* (BROYLES *et al.* 1991), přičemž tento guanosinový zbytek je konzervovaný i v rámci sekvence ATNTGA motivu UCS kvasinkových cytoplasmatických plasmidů. Stejně experimenty ukazují interakci VETF s několika adenosinovými zbytky UCE elementu *in vitro*, patrně v malém žlábkku DNA, což by vysvětlovalo VETF faktorem tolerovanou variaci nukleotidů A a T na ostatních pozicích UCE elementu (BROYLES *et al.* 1991; BROYLES 2003). Tyto výsledky tak naznačují možnou funkčnost promotorů pGKL plasmidů jako promotorů časných genů VACV po umístění do virového genomu, ačkoliv časně promotory VACV nefungují jako promotory pGKL plasmidů *in vivo*, pravděpodobně díky již zmíněné adaptaci transkripčního systému pGKL plasmidů na hexanukleotidový konsenzus motiv promotoru, jehož sekvence je patrně pro funkci *in vivo* invariantní.

### 6.1.8 Terminace transkripce a složení 3' konců mRNA pGKL plasmidů

Můj školitel zmapoval 3' konce mRNA všech genů pGKL plasmidů za použití 3' RACE-PCR experimentů. Bioinformatickou analýzou se mi podařilo identifikovat 1 až 4 předpokládané RNA vlásenky blízko detekovaných 3' konců mRNA každého ORF. Předpokládané struktury RNA vlásenek vykazovaly poměrně vysoké hodnoty Gibbsovy volné energie (medián  $-7,5$  kcal/mol) ve srovnání s predikovanými RNA vlásenkami vnitřních terminátorů transkripce modelových bakterií *E. coli* (vrchol rozdělení  $-14$  kcal/mol) a *Bacillus subtilis* (vrchol rozdělení  $-16$  kcal/mol) (DE HOON *et al.* 2005). Hodnoty Gibbsovy volné energie předpokládaných RNA vlásenek transkriptů pGKL plasmidů jsou ovšem srovnatelné s předpokládanými vnitřními terminátory bakterie *Mycoplasma hyopneumoniae* (medián  $-8,0$  kcal/mol), což je organismus se srovnatelně vysokým obsahem AT bází, jako kvasinkové lineární plasmidy (Tabulka 43). Předpokládané struktury RNA vlásenek pGKL plasmidů vykazovaly značnou heterogenitu v délce stonku (medián 15 bp), délce smyčky (medián 7 nt) a celkovém tvaru, přičemž stonky RNA vlásenek zpravidla nejsou plně dvouvláknové (medián 5 nezpárovaných nt) a nevykazují zvýšený výskyt GC bází (data nezobrazena). Tím se struktury RNA vlásenek pGKL plasmidů liší od kanonických GC bohatých RNA vlásenek vnitřních terminátorů transkripce bakterie *E. coli*, jejichž stonky mají průměrnou délku 8 bp a smyčka má průměrnou délku 4 nt (D'AUBENTON

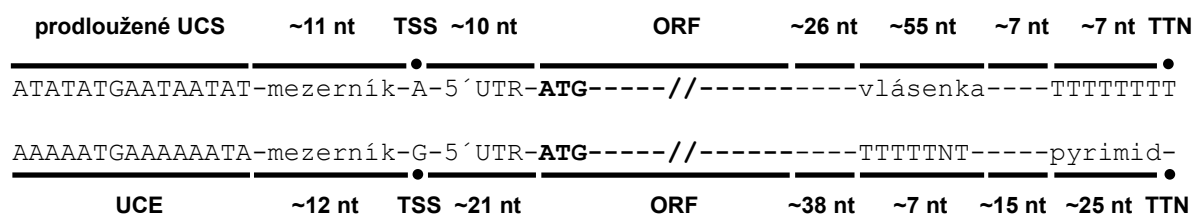


CARAFI *et al.* 1990; DE HOON *et al.* 2005). Nové bioinformatické přístupy však ukázaly, že celková délka a tvar RNA vlásenek vnitřních terminátorů transkripce je mezi jednotlivými bakteriálními druhy, a do jisté míry i mezi jednotlivými geny daného druhu, značně variabilní (MITRA *et al.* 2009).

**Tabulka 43:** Porovnání vlastností předpokládaných vnitřních terminátorů transkripce u AT bohatých genomů

Genom	Obsah AT (%)	Počet terminátorů	Medián $\Delta G$ (kcal/mol)	Medián vzdálenosti RNA vlásenky od stop kodónu (nt)	Medián celkové délky RNA vlásenky (nt)	Reference
<i>Mycoplasma hyopneumoniae</i>	71,5	558	-8,0	60,5	44	FRITSCH <i>et al.</i> (2015)
pGKL plasmidy	74,2	24	-7,5	26,0	55	Tato práce

Struktury RNA vlásenek vnitřních terminátorů transkripce u bakterií jsou obvykle následovány 7–8 nt dlouhým oligo(U) traktem, který podporuje zastavení RNAP na slabé DNA:RNA hybridní molekule obsahující dA-rU nukleotidové páry (MARTIN a TINOCO 1980; D'AUBENTON CARAFI *et al.* 1990; GUSAROV a NUDLER 1999). Naše původní analýzy přítomnost oligo(U) traktu po strukturách předpokládaných RNA vlásenek pGKL plasmidů neprokázaly. Jelikož je známo, že oligo(U) konec není nezbytně nutný pro účinnou vnitřní terminaci transkripce u bakterií jak *in vitro*, tak *in vivo* (INGHAM *et al.* 1995; UNNIRAMAN *et al.* 2001), předpokládali jsme, že AU bohaté sekvence transkriptů následující po RNA vlásenkách lineárních plasmidů by mohly mít podobný účinek na pozastavení RNAP a tání DNA:RNA hybridní molekuly během terminace transkripce jako kanonický U-konec. Detailnější sekvenční analýza však ukázala nabohacení thymidinových nukleotidů na 8 terminálních nukleotidech 3' konců cDNA genů pGKL plasmidů, což odpovídá předpokládanému oligo(U) traktu. Důležité je, že se podařilo detekovat přímou spojitost mezi předpokládanou strukturou RNA vlásenek a definovanou tvorbou 3' konců transkriptů pGKL plasmidů *in vivo*, což naznačuje mechanismus vnitřní terminace transkripce kvasinkových lineárních plasmidů, obdobný jako u bakterií [pro přehled viz (RAY-SONI *et al.* 2016)]. Přestože jsme neanalyzovali efektivitu terminace transkripce, předpokládáme, že transkripce se běžně pročítá alespoň za některé z předpokládaných terminátorových sekvencí transkripce. Jinak by nebyla možná funkční exprese přibližně poloviny otevřených čtecích rámců kvůli kompaktní organizaci genomu pGKL plasmidů. Struktury RNA vlásenek pGKL plasmidů budou patrně samostatnými funkčními prvky, protože definovaná tvorba 3' konců plasmidových transkriptů nebyla ovlivněna promotorem ani délkou či sekvencí kódující oblasti genu *in vivo*. Schématické zobrazení umístění sekvenčních a strukturních motivů spojených s místy iniciace a terminace transkripce pGKL plasmidů na základě této práce je na Obrázku 75.



**Obrázek 75:** Prvky spojené s místy iniciace (TSS, černá tečka) a terminace (TTN, černá tečka) transkripce pGKL plasmidů (horní část obrázku) a časných genů viru *Vaccinia* (spodní část obrázku). Zobrazeno je kódující vlákno DNA a nejčastěji se objevující nukleotid na každé pozici prodlouženého UCS a UCE motivu. Pro každý označený prvek je zobrazen medián jeho délky na základě sekvencí 5' a 3' konců mRNA pGKL plasmidů, získaných pomocí RACE-PCR experimentů V. Vopálenským, a 5' a 3' konců mRNA časných genů viru *Vaccinia*, získaných pomocí celogenomového sekvenování transkriptomu VACV (YANG *et al.* 2011a). Při syntéze porovnávaných transkriptů patrně dochází k obdobným mechanismům iniciace transkripce, ale rozdílným mechanismům terminace transkripce. Porovnávané transkripty mají podobný medián délky 5' a 3' nepřekládaných oblastí. Iniciátorový region promotorů pGKL plasmidů není zobrazen. Transkripty časných genů viru *Vaccinia* navíc obsahují 3' poly(A) konec.

### 6.1.9 Předpokládané viru podobné částice pGKL plasmidů

Geny *ORF9* kvasinkových cytoplasmatických lineárních plasmidů kódují podle naší bioinformatické analýzy předpokládaný homolog *scaffold* proteinu D13 viru *Vaccinia*, jenž se podílí na formaci nezralých sférických částic VACV (SZAJNER *et al.* 2005). Předpokládané terciární struktury plasmidových proteinů K2ORF9p a pSKL-ORF9p však postrádají 75 aminokyselinových zbytků N-konce proteinu D13, přičemž na základě dostupné literatury je možné, že právě N-konec proteinu D13 je důležitý z hlediska regulace procesu zrání částic VACV zprostředkované pomocí dalších několika virových proteinů (BISHT *et al.* 2009; BAHAR *et al.* 2011; UNGER *et al.* 2013). Uvážíme-li obdobnou biologickou roli pro proteiny ORF9p jako pro protein D13, tj. formaci membránových struktur kolem genomu daného genetického elementu, má absence předpokládané původní N-koncové regulační oblasti proteinu ORF9p smysl, neboť umožňuje menší celkovou kódující kapacitu genetického elementu, protože geny kódující regulační proteiny ovlivňující zrání částic budou nadbytečnými. Navíc, pokud pGKL plasmidy tvoří částice podobné nezralým částicím VACV, další maturace částic a obalení lipidovými dvojvrstvami v hostitelském prostředí kvasinkových buněk obsahujících buněčnou stěnu by pravděpodobně nemohla vést k produkci infekčních extracelulárních virových částic.

Pokud pGKL plasmidy v kvasinkových buňkách tvoří nezralé částice, bude tak jejich biologie pravděpodobně připomínat další poměrně rozšířenou skupinu cytoplasmatických kvasinkových genetických elementů, kterými jsou dvouvláknové RNA *killer* viry čeledi *Totiviridae*, jenž v hostitelských buňkách tvoří intracelulární neinfekční částice s chybějící vnější vrstvou virionu typickou pro viry využívající primárně extracelulární transmissi [pro přehled viz

(WICKNER *et al.* 2013)]. Předpokládané viru podobné částice pGKL plasmidů se mi však nepodařilo izolovat ultracentrifugací. Nezralé částice VACV, které by měly být stukturně příbuzné předpokládaným částicím lineárních plasmidů, nebyly dosud rozsáhle studovány. Lipidové složení nezralých částic VACV je neznámé a proteinové složení je částečně známé díky experimentům využívajícím imunoelektronovou mikroskopii [pro přehled viz (LALIBERTE a MOSS 2010)]. Při izolaci VACV využívající ultracentrifugaci v sacharosovém gradientu jsou nezralé částice VACV nestabilní (VANSLYKE *et al.* 1993). Pouze v jednom případě se podařilo pomocí ultracentrifugace v sacharosovém gradientu izolovat částice pravděpodobně odpovídající nezralým částicím VACV (SAROV a JOKLIK 1973), přičemž pro stabilitu částic mohla být důležitá přítomnost 0,5 % (w/v) neiontového detergentu Brij<sup>®</sup> 58 ve vzorku. Stabilita předpokládaných částic lineárních kvasinkových plasmidů při izolaci by tak mohla být teoreticky závislá na přítomnosti neiontového detergentu, což je ve shodě se zvýšenou stabilitou DNA pGKL plasmidů izolovaných z protoplastů proti působení deoxyribonukleáz za přítomnosti 0,5–1% (v/v) neiontového detergentu Triton X-100 (M. Pospíšek, osobní sdělení).

Pokud viru podobné částice pGKL plasmidů existují, nejsou příliš stabilní ve vysokých koncentracích chloridu cesného, neboť při izolaci využívající lyzát kvasinkových protoplastů a ultracentrifugaci v gradientu chloridu cesného byla naprostá většina DNA pGKL plasmidů téměř úplně oddělená od proteinů. Navíc není možné tyto pravděpodobné viru podobné částice efektivně uvolnit z buněk přípravou lyzátu pomocí skleněných kuliček a homogenizátoru, neboť při izolaci využívající tento druh lyzátu zůstala většina DNA pGKL plasmidů v peletu obsahujícím buněčné zbytky (data nezobrazena). V lyzátu se poté vyskytovalo poměrně malé množství pGKL plasmidů, což ztěžovalo následnou detekci DNA pGKL plasmidů po ultracentrifugaci v gradientu sacharózy pomocí elektroforetické analýzy. Pro jakékoliv další pokusy o izolaci předpokládaných viru podobných částic pGKL plasmidů ultracentrifugací by tak možná bylo výhodnější použít některý z připravených kmenů *K. lactis* s proteiny pGKL plasmidů označenými značkami, jejichž přítomnost v jednotlivých frakcích by měla být pomocí metody *western blot* snadno detekovatelná.

### 6.1.10 Evoluční původ kvasinkových lineárních plasmidů

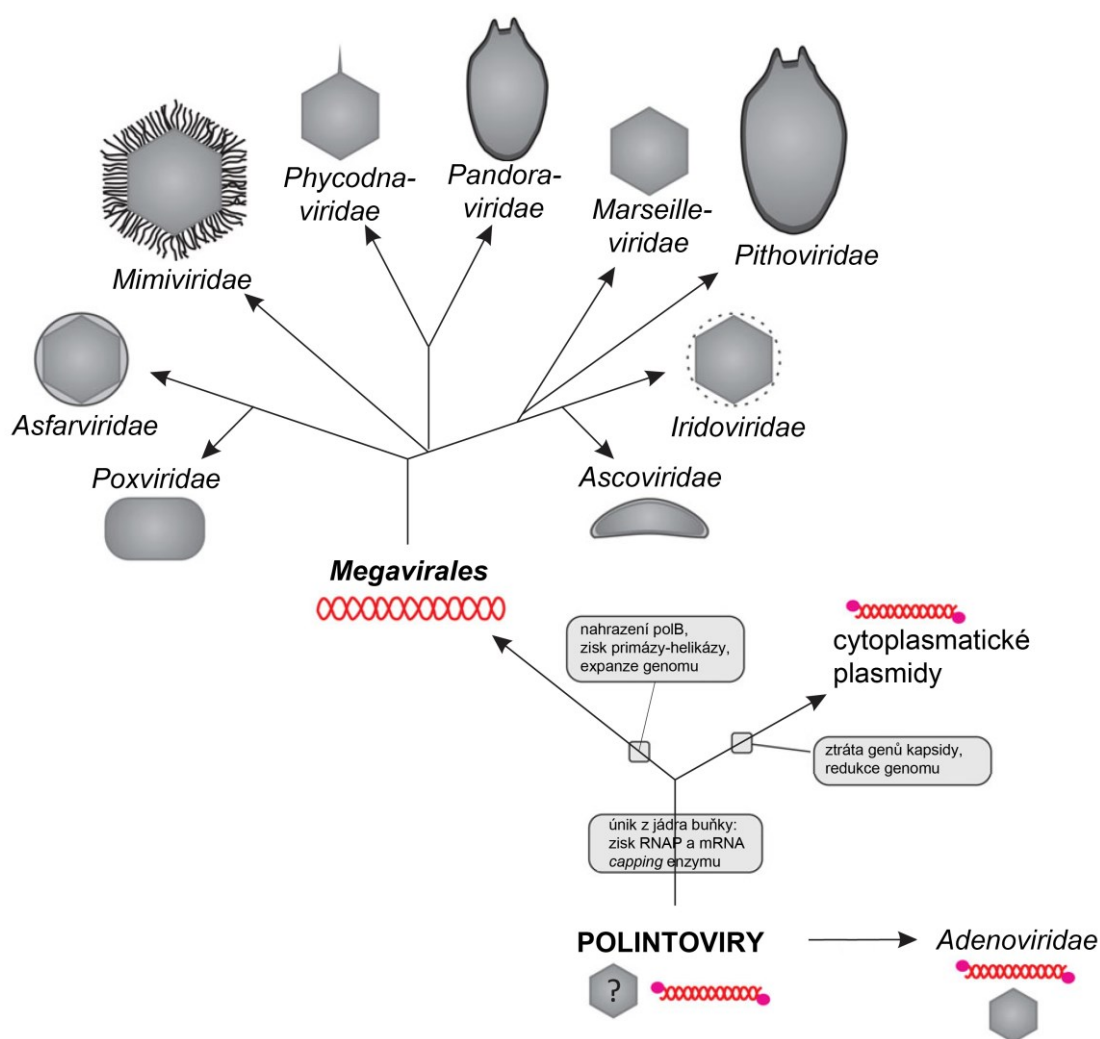
Určení evolučního původu kvasinkových cytoplasmatických plasmidů je poměrně náročné, protože otevřené čtecí rámce kódované těmito plasmidy jsou značně odlišné od dostupných sekvencí otevřených čtecích rámců v biologických databázích, a to jak na úrovni nukleotidů, tak na úrovni aminokyselinových zbytků. Z toho důvodu je kromě obtížnosti hledání homologních genů i velmi obtížné určit přibližné časové období, ve kterém došlo k oddělení

homologních genů od posledního společného předka. Proto se názor na evoluční původ těchto elementů postupem času pozvolně měnil, přičemž recentně se původ těchto elementů, stejně jako původ nukleocytoplasmatických virů, opět dostává do centra zájmu evolučních biologů.

Pro původní fylogenetické analýzy lineárních plasmidů byly použity geny kódující terminální protein fúzovaný s DNA polymerázou (KEMPKE *et al.* 1992; ROHE *et al.* 1992). Replikaci využívající DNA polymerázu a terminální protein používají kromě cytoplasmatických lineárních plasmidů některé lineární bakteriofágy, adenoviry a mitochondriální lineární plasmidy rostlin a hub. Původní analýzy shodně detekovaly blízkost DNA polymeráz mitochondriálních plasmidů k DNA polymerázám bakteriofágů typu  $\phi 29$ , přičemž DNA polymerázy cytoplasmatických plasmidů jsou od této skupiny zřetelně odděleny a mají původ v DNA polymeráze virů čeledi *Adenoviridae* či v DNA polymeráze společného předka cytoplasmatických plasmidů a adenovirů (KEMPKE *et al.* 1992; ROHE *et al.* 1992). Detekovaný odlišný fylogenetický původ DNA polymeráz mitochondriálních plasmidů rostlin a hub od DNA polymeráz cytoplasmatických plasmidů kvasinek je zajímavý, neboť pouze u těchto dvou skupin dědičných elementů se vyskytuje terminální protein fúzovaný s DNA polymerázou. Na společný původ mitochondriálních a cytoplasmatických plasmidů by mohlo poukazovat i nalezení mitochondriálních lineárních plasmidů u některých druhů kvasinek rodu *Pichia* (BLAISONNEAU *et al.* 1999). Tudíž by se teoreticky u kvasinek mohly vyskytovat mitochondriální a cytoplasmatické plasmidy současně, nicméně DNA polymerázy těchto kvasinkových mitochondriálních plasmidů mají blíže k DNA polymerázám mitochondriálních plasmidů jiných eukaryot než k DNA polymerázám cytoplasmatických plasmidů kvasinek (KLASSEN a MEINHARDT 2007).

Původní předpoklady o vzniku kvasinkových lineárních plasmidů z viru příbuzného virům čeledi *Adenoviridae* poněkud narušila sekvenční a posléze i fylogenetická analýza aminokyselinových zbytků otevřených čtecích rámců mRNA *capping* enzymů a předpokládaných helikáz lineárních plasmidů, která shodně poukazuje na příbuznost těchto otevřených čtecích rámců s proteiny nukleocytoplasmatických virů (BAYLIS *et al.* 1993; YÁÑEZ *et al.* 1993; LU *et al.* 1996; LARSEN *et al.* 1998; JESKE *et al.* 2007; SATWIKA *et al.* 2012). Díky tomu se názory na evoluční původ lineárních cytoplasmatických plasmidů začaly rozcházet. Původní a zároveň konzervativní názor podporuje vznik lineárních plasmidů z adenovirů (KLASSEN a MEINHARDT 2007). Druhý názor předpokládá složitější evoluční historii lineárních plasmidů, kdy některé geny pochází od adenovirů a některé od nukleocytoplasmatických virů, dnes souhrně klasifikovaných do patrně monofyletické virové skupiny *Megavirales* (IYER *et al.* 2001). Třetí názor se původem cytoplasmatických plasmidů nezabývá, ale nabízí variantu vzniku eukaryotických virů skupiny *Megavirales* právě z lineárních cytoplasmatických plasmidů (SHUMAN 2002), což by vysvětlilo fylogenetickou příbuznost mRNA *capping* enzymů a předpokládaných helikáz těchto elementů.

Recentní představy o původu kvasinkových cytoplasmatických lineárních plasmidů vycházejí z několika nových poznatků. Prvním z těchto poznatků je, že některé bakteriofágy, využívající k replikaci strategii bakteriofága typu  $\varphi 29$ , viry skupiny *Adenoviridae* a viry skupiny *Megavirales* kódují kapsidový protein (*scaffold* protein u poxvirů) ze stejné proteinové rodiny, což teoreticky naznačuje shodný evoluční původ virů skupin *Adenoviridae* a *Megavirales* (NANDHAGOPAL *et al.* 2002; HYUN *et al.* 2007; BAHAR *et al.* 2011; HYUN *et al.* 2011). Druhým z těchto poznatků je identifikace nové skupiny autonomně se replikujících eukaryotických DNA transpozonů pojmenovaných jako *Polinton* či *Maverick*, které také kódují terminální protein pravděpodobně fúzovaný s DNA polymerázou (KAPITONOV a JURKA 2006; PRITHAM *et al.* 2007). Fylogenetická analýza aminokyselinových sekvencí DNA polymeráz naznačila příbuznost DNA polymeráz polintonů s DNA polymerázami kvasinkových cytoplasmatických plasmidů (KAPITONOV a JURKA 2006; YUTIN *et al.* 2013), přičemž se dnes předpokládá, že polintony by mohly být jedním z chybějících článků propojujících pravděpodobně evolučně příbuzné viry skupin *Adenoviridae*, *Megavirales* a cytoplasmatické kvasinkové plasmidy (KRUPOVIČ a KOONIN 2015). Metagenomické analýzy ukázaly, že velké množství polintonů pravděpodobně kóduje kapsidový protein ze stejné proteinové rodiny jako kapsidový protein virů *Adenoviridae* a *Megavirales*, díky čemuž se předpokládá, že polintony by ve skutečnosti mohly být novou skupinou eukaryotických jaderných DNA virů s prozatímním názvem Polintoviry (KRUPOVIČ *et al.* 2014; YUTIN *et al.* 2015). Recentní evoluční pohled naznačuje, že z polintovirů se oddělily adenoviry, a z polintoviru, který unikl z jádra do cytoplasmy a přibral geny umožňující biosyntézu mRNA molekul, se stal společný předek kvasinkových cytoplasmatických plasmidů a virů skupiny *Megavirales* (KOONIN *et al.* 2015a). V dalším vývoji pak zřejmě došlo především ke ztrátě kapsidového proteinu u cytoplasmatických plasmidů a nárůstu genomu a nahrazení genu fúzního terminálního proteinu a DNA polymerázy u *Megavirales* (Obrázek 76).

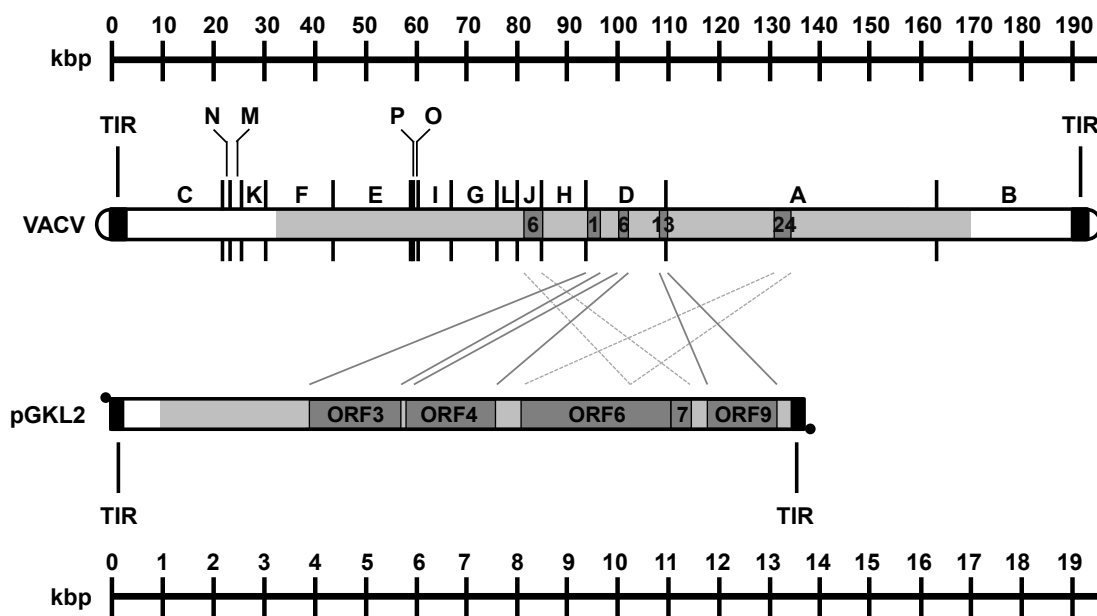


**Obrázek 76:** Předpokládaný vývoj velkých dvouvláknových DNA virů a cytoplasmatických plasmidů eukaryot z polintovirů. Pro každou rodinu velkých eukaryotických DNA virů je znázorněno zjednodušené schéma struktury virionu (šedě). Nukleová kyselina (červeně) a terminální proteiny (růžově) genomu jsou znázorněny. RNAP: DNA-dependentní RNA polymeráza, polB: DNA-dependentní DNA polymeráza rodiny B. Upraveno podle (KOONIN *et al.* 2015b).

Výsledky této disertační práce tak částečně zapadají do současného pohledu na evoluční vznik kvasinkových cytoplasmatických plasmidů, neboť pokud mají tyto lineární plasmidy společný původ s *Megavirales*, vysvětluje se tím mnou detekovaná fylogenetická příbuznost RNA polymeráz lineárních plasmidů s RNA polymerázou virů čeledi *Poxviridae*, a také dříve detekovaná fylogenetická příbuznost mRNA *capping* enzymu a předpokládané helikázy lineárních plasmidů s příslušnými proteiny nukleocytoplasmatických virů. Do současného pohledu na evoluční vznik kvasinkových cytoplasmatických plasmidů však poněkud nezapadá fakt, že jsem u lineárních plasmidů detekoval předpokládaný homolog *scaffold* proteinu D13 viru *Vaccinia*, přičemž homology tohoto proteinu tvoří hlavní kapsidový protein u *Adenoviridae*, všech čeledí skupiny *Megavirales* až na poxviry, a zřejmě i u předpokládaných polintovirů. To znamená, že kvasinkové cytoplasmatické plasmidy gen předpokládané kapsidy navzdory obecnému mínění



neztratily. Vzhledem k tomu, že protein K2ORF9p bude pravděpodobně strukturně nejvíc podobný proteinu D13 poxvirů a ne poněkud složitějšímu kapsidovému proteinu adenovirů či naopak jednoduššímu kapsidovému proteinu polintovirů, předpokládám, že nejbližšími, dosud charakterizovanými, příbuznými genetickými elementy cytoplasmatických plasmidů jsou právě viry z čeledi *Poxviridae*. Schematické znázornění umístění fylogeneticky příbuzných genů pGKL plasmidů a viru *Vaccinia* na základě této práce je na Obrázku 77.



**Obrázek 77:** Organizace genomu viru *Vaccinia* kmene Copenhagen a autonomního plasmidu pGKL2. Geny konzervované u orthopoxvirů a autonomních cytoplasmatických kvasinkových plasmidů jsou zobrazeny světle šedě, oblasti obsahující variabilní geny na koncích genomů jsou zobrazeny bíle. Terminální invertované repetice (TIR) jsou zobrazeny jako černé obdélníky, terminální smyčky jsou zobrazeny jako černé půlelipsy, terminální proteiny jsou zobrazeny jako černé puntíky. Homologní geny jsou zobrazeny jako tmavě šedé obdélníky. U každého genomu je uvedeno měřítko v tisících párů bází. Geny viru *Vaccinia* nesou písmenné označení (A–O), odpovídající jednotlivým DNA fragmentům genomu viru kmene Copenhagen, vzniklým restrikcčním štěpením endonukleázou *HindIII*.

Domněnku o společném původu cytoplasmatických plasmidů a poxvirů dále podporují i další výsledky této disertační práce, které kromě fylogenetické analýzy aminokyselinových sekvencí proteinů poukazují i na pravděpodobný společný původ nukleotidových sekvencí promotorových elementů pGKL plasmidů a VACV a na celkovou vzájemnou podobnost mRNA pGKL plasmidů a VACV indikující, že při transkripci a translaci budou lineární plasmidy a poxviry v mnoha ohledech využívat obdobné strategie.

Při vzájemném porovnání homologních proteinů lineárních plasmidů a poxvirů je patrné, že proteiny lineárních plasmidů mají ve všech případech kratší délku polypeptidového řetězce a nižší molekulovou hmotnost, než jejich poxvirové protějšky. Nabízí se tak otázka, jestli

proteiny u společného předka lineárních plasmidů a poxvirů měly jednodušší architekturu podobnou proteinům lineárních plasmidů nebo naopak složitější architekturu podobnou proteinům poxvirů. V případě mRNA *capping* enzymu, předpokládané helikázy a předpokládaného *scaffold* proteinu nelze jednoznačně určit, která architektura byla původnější. V případě RNA polymerázy, kterou měl původní virový předek patrně získat od hostitele (KRUPOVIČ a KOONIN 2015), se však domnívám, že unikátní RNA polymeráza lineárních plasmidů prošla rozsáhlou redukcí původně složitého vícepodjednotkového enzymu. Při vývoji lineárních plasmidů ze společného předka tak zřejmě došlo k poměrně značné redukci velikosti a celkového množství genů vedoucí k maximální kompaktizaci genomu.

### 6.1.11 Translace mRNA pGKL plasmidů

O mechanismu translace lineárních plasmidů není mnoho známo. Díky cytoplasmatické lokalizaci pGKL plasmidů se nejdříve předpokládalo, že by jejich mRNA mohly být překládány translačním aparátem mitochondrií. Tato možnost však byla záhy vyloučena, neboť k expresi mRNA pGKL plasmidů dochází i v buňkách postrádajících mitochondriální DNA. Navíc sekvenované proteinové produkty pGKL plasmidů dokládají použití univerzálního genetického kódu bez speciálních výjimek mitochondriálního genetického kódu při translaci plasmidových mRNA (STARK *et al.* 1990). Translace mRNA pGKL plasmidů buněčným translačním aparátem bude patrně negativně ovlivněna vysokým obsahem AT bází plasmidových genů, díky čemuž je využívání kodónů plasmidových ORF odlišné od kodónů používaných u jaderných kvasinkových genů vykazujících vysokou míru exprese (STARK *et al.* 1990). Ačkoliv výzkum translace lineárních plasmidů nebyl primární náplní této disertační práce, některé výsledky této práce spolu s výsledky mého školitele umožňují diskutovat aspekty iniciace translace mRNA pGKL plasmidů.

Výsledky z kmene nesoucího bodové mutace v aktivním centru methyltransferázové domény K2ORF3p naznačují, že přítomnost čepičky typu *cap* 0 na transkriptech lineárních plasmidů je pro plasmidy esenciální. To je v kontrastu se zjištěním, že plasmidové mRNA kódující *killer* toxin neinteragují s eIF4E faktorem vázajícím čepičku *in vitro* a že translace plasmidových mRNA kódujících toxin zřejmě není na translačním faktoru eIF4E závislá *in vivo* (VOPÁLENSKÝ *et al.* 2018). Mnou provedená statistická analýza ukázala na signifikantní negativní korelaci mezi délkou 5' poly(A) leaderu a přítomností 5' čepičky na transkriptech plasmidových genů. Díky tomu je frekvence výskytu čepičkové struktury u transkriptů genů pGKL plasmidů vznikajících z promotoru s iniciátorovým regionem obsahujícím tři a více následných adenosinových zbytků poměrně nízká (22 %) oproti transkriptům genů pGKL plasmidů

vznikajících z promotoru obsahujícího jeden či dva následné adenosinové zbytky v místě iniciace transkripce (84 %). U této druhé skupiny genů, u jejichž transkriptů nedochází k efektivní netemplátové polyadenylaci, se navíc občas vyskytuje i poměrně dlouhá 5' nepřekládaná oblast (medián délky 60 nt) oproti genům, jejichž transkripty obsahují 5' poly(A) leader, kde je 5' nepřekládaná oblast (medián délky 9 nt) zpravidla tvořena pouze adenosinovými zbytky. Proto předpokládám, že možná pouze u této druhé a menší skupiny genů (*K2ORF2*, *K2ORF3* a *K2ORF8*) je přítomnost 5' čepičkové struktury esenciální pro efektivní míru translace jejich mRNA mechanismem využívajícím iniciaci translace závislou na faktoru eIF4E. Tento předpoklad je v souladu se zjištěním, že efektivní exprese genů toxinu produkujících 5' polyadenylované transkripty přítomnost faktoru eIF4E *in vivo* nevyžaduje (VOPÁLENSKÝ *et al.* 2018), čili mechanismus iniciace translace plasmidových transkriptů s 5' poly(A) leaderem bude zřejmě na čepičkové struktuře mRNA nezávislý.

Pro rozdílné mechanismy iniciace translace plasmidových mRNA s a bez 5' poly(A) leaderu by mohly hovořit i výsledky z kmene nesoucího mutace rušící následné adenosinové nukleotidy v iniciátorovém regionu promotoru  $\gamma$  podjednotky toxinu, který vykazoval sníženou míru exprese *killer* toxinu *in vivo*. Obdobné mutace iniciátorového regionu promotorů středních a pozdních genů VACV také vedou ke snížené míře exprese reportérového genu *in vivo*, ačkoliv tyto výsledky jsou interpretovány předpokládanou sníženou silou promotoru, přestože hladina RNA v těchto experimentech nebyla zkoumána (DAVISON a MOSS 1989b; BALDICK *et al.* 1992; KNUTSON *et al.* 2006). Nelze proto vyloučit, že rozdílná míra exprese reportérového genu z mutovaných promotorů VACV a pGKL plasmidů by mohla být ovlivněna na úrovni iniciace translace. Pro nižší míru translace plasmidových mRNA neobsahujících 5' poly(A) leader by nepřímo mohl hovořit i fakt, že 5' nepolyadenylované transkripty genu *K2ORF3* vznikají *in vivo* ze dvou promotorů. Moje výsledky dokládají, že proteiny K2ORF3p, K2ORF6p a K2ORF7p tvoří, pravděpodobně v ekvimolárním poměru, transkripční komplex pGKL plasmidů *in vivo*. Exprese K2ORF3p ze dvou promotorů by tak zřejmě přispívala k vyrovnání buněčné hladiny proteinu K2ORF3p s hladinou proteinů K2ORF6p a K2ORF7p překládaných z 5' polyadenylovaných transkriptů. Částečně ve shodě s mým předpokladem nižší míry translace plasmidových mRNA postrádajících 5' poly(A) leader by mohly i výsledky SCHICKEL *et al.* (1996), které z testovaných promotorů detekovaly nejnižší míru exprese reportérového genu z pGKL plasmidu při použití promotoru genu *K2ORF2*.

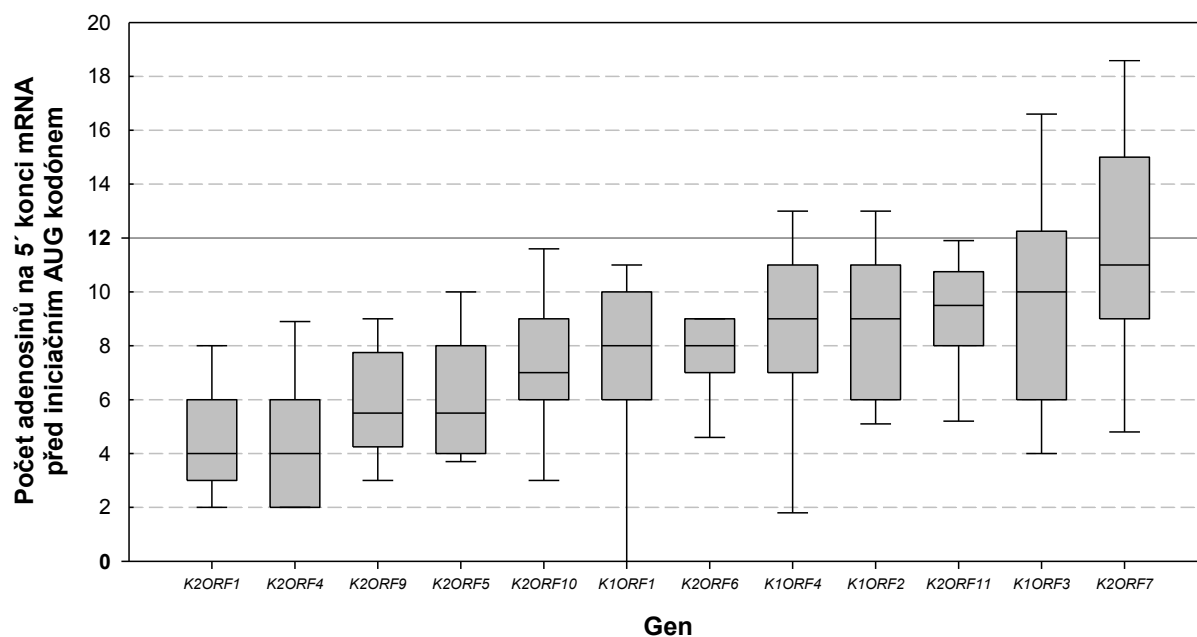
Pro pochopení možných mechanismů uplatňujících se při iniciaci translace plasmidových mRNA s 5' poly(A) leaderem jsou velmi důležité poznatky získané studiemi zabývajícími se mechanismy translace transkriptů VACV. Původní studie zabývající se porovnáním translace buněčných a virových mRNA v *in vitro* translačním systému detekovala rozdílnou citlivost

transkriptů HeLa buněk a časných a postreplikativních transkriptů VACV na inhibici translace, indukovanou analogem mRNA čepičky a polyadenylovanými RNA molekulami různých délek, a rozdílné požadavky na translační iniciační faktory pro reverzi inhibice translace (BABLANIAN *et al.* 1991). Obecně vykazovaly virové mRNA menší míru inhibice translace indukované analogem mRNA čepičky a 3' polyadenylovanými RNA (BABLANIAN *et al.* 1991). Předpokládanou menší závislost translace virových mRNA na přítomnosti mRNA čepičky podporuje i termosenzitivní mutanta VACV s mutacemi v guanylyltransferázové a methyltransferázové doméně mRNA *capping* enzymu D1, kdy mutanta při permissivních i nepermissivních podmínkách vykazuje srovnatelnou hladinu syntézy virových mRNA a proteinů (HASSETT *et al.* 1997). Virus *Vaccinia* vykazuje *in vivo* malou závislost na intaktním komplexu vázajícím mRNA čepičku zvaném eukaryotický iniciační faktor 4F (eIF4F; heterotrimer proteinů eIF4E, eIF4G a eIF4A), neboť v buňkách s rozštěpeným faktorem eIF4G přetrvává syntéza pozdních virových proteinů (MULDER *et al.* 1998).

Další pokusy demonstrovaly, že reportérové mRNA postrádající čepičku, ale obsahující 5' poly(A) leadery o délce 5, 12 a 25 nukleotidů, jsou se vzrůstající efektivitou účinně translatovány v eukaryotických bezbuněčných systémech *in vitro* (GUDKOV *et al.* 2005). Následná studie zabývající se enhancerovým efektem 5' poly(A) leaderů na translaci ukázala, že na reportérové mRNA obsahující 25 nt 5' poly(A) leader a postrádající čepičku dochází k tvorbě 48S iniciačních ribozomálních komplexů na AUG kodónu i bez přítomnosti eukaryotického iniciačního faktoru 3 (eIF3) a faktoru eIF4F *in vitro* (SHIROKIKH a SPIRIN 2008). Tvorba 48S iniciačních ribozomálních komplexů na mRNA s 5' poly(A) leaderem dokonce předčí tvorbu 48S iniciačních ribozomálních komplexů na mRNA s 5' čepičkou za přítomnosti faktorů eIF3 a eIF4F v přímé kompetici *in vitro* (SHIROKIKH a SPIRIN 2008). Detailní analýza sekvencí 5' nepřekládaných oblastí transkriptů *S. cerevisiae* také poukázala na enhancerový efekt 5' poly(A) sekvencí na translaci, neboť mRNA obsahující do 11 následných adenosinů před AUG kodónem jsou translatovány se vzrůstající mírou *in vivo*, čemuž odpovídá i vysoká hladina proteinů kódovaných těmito mRNA *in vivo* (XIA *et al.* 2011). Naopak mRNA *S. cerevisiae* obsahující 12 a více následných adenosinů před AUG kodónem vykazuje nejnižší míru translace a nejnižší denzitu ribozómů *in vivo*, čemuž odpovídá i nejnižší hladina proteinů kódovaných těmito mRNA *in vivo* (XIA *et al.* 2011). Předpokládá se, že inhibiční efekt translace těchto mRNA by mohl být způsoben vazbou 5' nepřekládaných oblastí PABP1 proteinem, neboť PABP1 protein má vysokou afinitu pro oligo(A) nukleotidy obsahující 12 a více adenosinů *in vitro* (SACHS *et al.* 1987; XIA *et al.* 2011). Bylo také ukázáno, že PABP1 negativně autoreguluje svoji syntézu vazbou A-bohaté 5' nepřekládané oblasti svojí mRNA *in vitro* a *in vivo* (BAG a WU 1996; PATEL *et al.* 2005).

Výsledky analýzy 5' nepřekládaných oblastí mRNA *S. cerevisiae* jsou ve značné shodě s recentní analýzou, která demonstruje, že reportérové mRNA obsahující 5' poly(A) leader o počtu 4–12 následných adenosinů před AUG kodónem jsou translatovány se vzrůstající mírou v HeLa buňkách infikovaných VACV, což těmto mRNA poskytuje *in vivo* translační výhodu nad transkripty bez 5' poly(A) leaderu v postreplikativním stádiu infekce VACV (DHUNGEL *et al.* 2017). Reportérové mRNA jsou v buňkách infikovaných VACV preferenčně translatovány, i když na svém 5' konci obsahují analog mRNA čepičky, který nemůže být vázán translačním faktorem eIF4E, a i v případech, kdy je množství dostupného eIF4E v buňce sníženo pomocí malé interferující RNA či nízkomolekulárního inhibitoru kinázy fosforylující vazebný protein eIF4E proteinu (DHUNGEL *et al.* 2017). V infikovaných buňkách tak 5' poly(A) leader poskytuje reportérové mRNA schopnost používat iniciaci translace nezávislou na čepičkové struktuře, ačkoliv nefunguje jako interní místo pro vstup ribozomu (DHUNGEL *et al.* 2017).

Analogicky tak výše uvedené výsledky v kombinaci s výsledky mého školitele naznačují, že by 5' poly(A) leader plasmidových mRNA mohl fungovat jako enhancer translace nezávislé na mRNA čepičce. Délka 5' poly(A) leaderu plasmidových mRNA se u jednotlivých genů pGKL plasmidů liší, přičemž moje výsledky indikují, že počet netemplátově zařazených adenosinů 5' poly(A) leaderu vzrůstá s vzrůstajícím počtem následných adenosinových zbytků kódujícího vlákna DNA v iniciátorovém regionu promotoru. Medián délky 5' poly(A) leaderu před AUG iniciačním kodónem u plasmidových mRNA pro jednotlivé geny nabývá hodnoty 4–11 (Obrázek 78), což jsou hodnoty, při kterých by měl 5' poly(A) leader mRNA u kvasinek fungovat jako translační enhancer (XIA *et al.* 2011). Zdá se, že by u většiny genů pGKL plasmidů mohlo skrze rozdílnou délku 5' poly(A) leaderu mRNA, a tudíž i předpokládanou rozdílnou míru translace mRNA, docházet k regulaci exprese zajišťující různé hladiny jednotlivých plasmidových proteinů *in vivo*. U transkriptů genů *K1ORF2*, *K1ORF3*, *K1ORF4* a *K2ORF7* byly běžně detekovány transkripty obsahující 12 a více adenosinů před AUG kodónem (Obrázek 78), což znamená, že translace určité frakce transkriptů těchto genů by mohla být negativně regulována PABP1 proteinem. To je ve shodě s detekovanou zvýšenou expresí *killer* toxinu pGKL plasmidů v kmeni *K. lactis* s deletovaným *PAB1* genem (VOPÁLENSKÝ *et al.* 2018).



**Obrázek 78:** Krabicový graf počtu adenosinů na 5' konci mRNA před AUG iniciačním kodómem translace u genů pGKL plasmidů, většina jejichž transkriptů obsahovala 5' poly(A) leader. Spodní a vrchní část boxu indikuje první a třetí kvartil. Úsečky indikují 10. a 90. percentil. Odlehle hodnoty nejsou vyznačeny. Medián je vyznačen čarou. Geny jsou seřazeny zleva doprava podle předpokládané vzrůstající míry translace jejich mRNA. Hodnota 12 adenosinů před AUG kodómem je vyznačena šedou nepřerušovanou čarou.

Enhancerový efekt 5' poly(A) sekvencí na translaci mRNA *S. cerevisiae*, mRNA VACV v infikovaných buňkách, a teoreticky i na translaci mRNA lineárních plasmidů, by mohl být limitován délkou 11–12 následných adenosinů před AUG kodómem ještě z jiného důvodu kromě možné vazby PABP1 proteinu. Nedávno bylo kombinací *in vitro* a *in vivo* přístupů u bakterií a kvasinek ukázáno, že ribozom obousměrně klouže na poly(A) sekvencích obsahujících 11 a více adenosinů (KOUTMOU *et al.* 2015), přičemž se předpokládá, že 5' poly(A) leadery delší než 11 adenosinů by měly negativní efekt na iniciaci translace mRNA snížením rozpoznání iniciačního AUG kodómu vlivem klouzání ribozomu (JHA *et al.* 2017; WALSH 2017). Podle literatury a mých výsledků jsou však 5' poly(A) leadery především u transkriptů pozdních genů VACV běžně delší než 11 adenosinů (YANG *et al.* 2012). Nová práce ukázala, že VACV v buňkách virovou kinázou B1 fosforyluje protein malé podjednotky ribozomu, který se nazývá receptor pro aktivovanou kinázu C 1 (RACK1). Tato modifikace hostitelského translačního aparátu poskytuje reportérové mRNA obsahující dlouhý 5' poly(A) leader o počtu 25 následných adenosinů před AUG kodómem translační výhodu (JHA *et al.* 2017). Předpokládá se, že fosforylace RACK1 snižuje rychlost skenování mRNA ribozomem, což má negativní vliv na translaci buněčných mRNA, která je závislá na čepičce, ale pozitivní vliv na translaci mRNA obsahující dlouhý 5' poly(A) leader, kde pomalá rychlost pohybu ribozomu po mRNA působí proti obousměrnému klouzání ribozomu na poly(A) sekvencích snižujícím rozpoznání



iniciačního kodónu (JHA *et al.* 2017; WALSH 2017). Ačkoliv kvasinkové lineární plasmidy ještě obsahují některé ORF se zatím neznámou funkcí, nepředpokládám, že by kódovaly homolog kinázy B1 VACV. I to může být jedním z důvodů, proč medián délky 5' poly(A) leaderu před AUG iniciačním kodónem u plasmidových mRNA nepřesahuje hodnotu 11. Negativní vliv příliš dlouhého 5' poly(A) leaderu na translaci plasmidových mRNA předpokládaným klouzáním ribozomu by šlo teoreticky testovat vnesením několika adenosinů do netemplátového vlákna za iniciátorový region plasmidového promotoru, který by byl použit k expresi reportérového genu, jež by byla sledována v kmeni s deletovaným *PAB1* genem.

## 6.2 POXVIRY

### 6.2.1 Složení 5' konců mRNA viru *Vaccinia*

Výsledky 5' RACE-PCR analýzy genů pGKL plasmidů mého školitele ukázaly, že 5' polyadenylované transkripty běžně neobsahují mRNA čepičkovou strukturu. Na získaných datech se mi dokonce podařilo detekovat signifikantní negativní korelaci mezi počtem netemplátových adenosinů 5' poly(A) leaderu a přítomností 5' mRNA čepičky. Poté, co moje výsledky ukázaly, že 5' poly(A) leader transkriptů pGKL plasmidů pravděpodobně vzniká obdobným mechanismem jako 5' poly(A) leader transkriptů VACV, začalo mě zajímat složení 5' konců mRNA VACV. Ačkoliv obecně převládá čtyři dekády trvající názor, že většina transkriptů časných a postreplikativních genů VACV bude obsahovat 5' mRNA čepičku (BOONE a MOSS 1977), na úrovni jednotlivých transkriptů tato premisa nikdy testována nebyla.

Moje výsledky 5' RACE-PCR analýz několika vybraných genů naznačují, že přítomnost 5' mRNA čepičky je u časových tříd genů VACV rozdílná, přičemž stejně jako u pGKL plasmidů přítomnost 5' čepičky na transkriptech VACV signifikantně negativně koreluje s rostoucím počtem netemplátových adenosinů 5' poly(A) leaderu mRNA. Jelikož délka 5' poly(A) leaderu mRNA u každé následné časové třídy genů VACV podle mých výsledků i podle literatury signifikantně vzrůstá (YANG *et al.* 2012), vysvětluje se tak detekovaná snižující se frekvence výskytu 5' mRNA čepičky u každé následné časové třídy genů VACV. Zmíněná negativní korelace poukazuje na možnost, že míra 5' polyadenylace mRNA nějakým dosud neznámým mechanismem přímo ovlivňuje syntézu čepičky na 5' koncích mRNA (či odstranění čepičky z 5' konců mRNA). Tuto možnost dále podporují moje výsledky ukazující, že na úrovni jednotlivých pozdních genů viru *Vaccinia* jsou transkripty obsahující mRNA čepičku na 5' koncích signifikantně kratší, než transkripty neobsahující 5' čepičkovou strukturu. Tyto výsledky také poukazují na možnost, že celková délka 5' poly(A) leaderu mRNA postreplikativních genů VACV vypočítaná na základě celogenomových analýz transkriptomu VACV může být podhodnocená, protože pro sekvenování byly použity pouze mRNA obsahující 5' čepičku (YANG *et al.* 2012). Zdá se, že moje výsledky tento předpoklad podporují, protože zatímco medián celkové délky 5' poly(A) leaderu mRNA u středních genů VACV je na základě literatury (8 nt) a mých výsledků (7 nt) obdobný, medián celkové délky 5' poly(A) leaderu mRNA u pozdních genů VACV je na základě literatury (11 nt) a mých výsledků (16 nt) odlišný. Publikovanému mediánu celkové délky 5' poly(A) leaderu mRNA pozdních genů VACV (11 nt) jsou bližší moje výsledky získané 5' RLM-RACE analýzou vybraných pozdních genů VACV (9 nt), která také vychází pouze z transkriptů obsahujících 5' čepičku.

Porovnání mých výsledků získaných 5' RACE-PCR a 5' RLM-RACE analýzou dvou vybraných pozdních genů VACV také umožňuje diskutovat literaturou popsany fenomén všudypřítomné iniciace transkripce VACV (YANG *et al.* 2012). Pomocí 5' RLM-RACE analýzy bylo pro geny *A17L* a *C3L* VACV detekováno více specifických produktů, zatímco pomocí klasické 5' RACE-PCR analýzy byl pro oba geny detekován pouze jeden specifický produkt. Protože analýzy vycházely ze stejného vzorku RNA a amplifikace cDNA probíhala za obdobných podmínek, lze z výsledků vyvodit závěr, že minoritní transkripční varianty pozdních genů VACV naproti majoritním transkripčním variantám převážně obsahují 5' čepičkovou strukturu a jejich četnost *in vivo* je ve srovnání s majoritními transkripčními variantami podstatně menší. K všudypřítomné iniciaci transkripce VACV často dochází uvnitř kódujících sekvencí otevřených čtecích rámců, přičemž v některých případech minoritní transkripční varianty kódují zkrácené isoformy proteinů a mohou pravděpodobně být i translatované (YANG *et al.* 2012; YANG *et al.* 2015).

Virus *Vaccinia* obsahuje dva geny (časný gen *D9R* a střední gen *D10R*), které kódují enzymy schopné odstraňovat methylovanou čepičkovou strukturu typu *cap 0* z 5' konců RNA *in vitro* (PARRISH a MOSS 2007; PARRISH *et al.* 2007). Nabízí se tak otázka, zda za detekovanou absenci 5' čepičkové struktury u většiny postreplicativních transkriptů může být zodpovědná aktivita těchto dvou virových enzymů *in vivo*. Transkripty časných genů obsahující i neobsahující 5' poly(A) leader spolu s transkripty středních genů VACV jsem analyzoval ze stejného výchozího vzorku získaného lyzí HeLa buněk 4 hod. po infekci virem, kde se teoreticky mohla projevit aktivita proteinu D9 i D10. Frekvence výskytu čepičkové struktury u transkriptů časných genů bez 5' poly(A) leaderu (86 %) byla srovnatelná s frekvencí výskytu čepičkové struktury u transkriptů genů pGKL plasmidů bez 5' poly(A) leaderu (87 %) a snižující se frekvence výskytu čepičkové struktury u časných a středních transkriptů VACV v závislosti na délce 5' poly(A) leaderu měla podobný trend, jako snižující se frekvence výskytu čepičkové struktury u transkriptů pGKL plasmidů v závislosti na délce 5' poly(A) leaderu. Protože pGKL plasmidy s největší pravděpodobností nekódují homolog proteinu D9 či D10, předpokládám, že za nízkou frekvenci výskytu čepičkové struktury u transkriptů časných a středních genů VACV s 5' poly(A) leaderem nebude primárně zodpovědná aktivita některého z virových enzymů odstraňujících 5' čepičku z mRNA. Nabízí se ovšem otázka, zda zmíněné enzymy neodstraňují čepičkovou strukturu z RNA s různou účinností v závislosti na přítomnosti či nepřítomnosti 5' poly(A) leaderu, čímž by k nízké frekvenci výskytu čepičkové struktury 5' polyadenylovaných transkriptů postreplicativních genů VACV alespoň přispívaly. V tomto ohledu však biochemické experimenty *in vitro* nasvědčují, že 5' čepička RNA molekuly připojená na 5' terminální adenosin by mohla být naopak více rezistentní k štěpení enzymem D10 (PARRISH *et al.* 2007).

### 6.2.2 Translace mRNA viru *Vaccinia*

Pochopení mechanismů uplatňujících se při iniciaci translace mRNA viru *Vaccinia* a analogický přenos těchto znalostí na pochopení translace mRNA pGKL plasmidů je poměrně náročnou záležitostí, protože VACV v buňkách rozsáhle mění složení a lokalizaci jednotlivých komponent translačního aparátu a navozuje globální potlačení syntézy proteinů hostitelské buňky, které se netýká pouze skupiny proteinů oxidativní fosforylace, jejichž produkce v infikovaných buňkách naopak vzrůstá [pro přehled například (LIEM a LIU 2016; CAO *et al.* 2017)].

Původní studie naznačila, že k potlačení syntézy buněčných proteinů v infikovaných buňkách by mohla přispívat selektivní translace mRNA viru, protože časné i postreplicative mRNA VACV vykazovaly při translaci *in vitro* sníženou závislost na některých translačních iniciačních faktorech oproti buněčným mRNA (BABLANIAN *et al.* 1991). Následné studie ukázaly, že oproti ostatním virům, jejichž mRNA jsou v buňkách preferenčně translatovány, VACV nesnižuje koncentraci proteinů iniciačního translačního komplexu vázajícího čepičku eIF4F a ani proteiny eIF4F komplexu nemodifikuje *in vivo* (GIERMAN *et al.* 1992; SCHNIERLE a MOSS 1992). Naopak VACV v buňkách podporuje formaci eIF4F komplexu stimulací signálních drah vedoucích k fosforylaci eIF4E proteinu a degradaci vazebného proteinu blokujícího funkci eIF4E proteinu, přičemž nízkomolekulární inhibitory těchto signálních drah vedou k potlačení replikace VACV a produkce virových proteinů (WALSH *et al.* 2008; ZABOROWSKA a WALSH 2009). Kromě podpory formace eIF4F komplexu VACV v buňce způsobuje redistribuci proteinů tohoto komplexu, kdy v infikovaných buňkách faktory eIF4G a eIF4E lokalizují do cytoplasmatických virových továren, kde dochází k replikaci virové DNA, přičemž změna lokalizace eIF4G faktoru vyžaduje interakci tohoto faktoru s virovým ssDNA vazebným proteinem I3 (KATSAFANAS a MOSS 2007; WALSH *et al.* 2008; ZABOROWSKA *et al.* 2012). Některé studie dokonce popsaly esencialitu intaktního eIF4F komplexu *in vivo* pro replikaci VACV v infikovaných buňkách (ALDABE *et al.* 1995; WELNOWSKA *et al.* 2009; MCMAHON *et al.* 2011). Přestože je podle těchto studií faktor eIF4G esenciální pro replikaci VACV, v infikovaných buňkách s rozštěpeným faktorem eIF4G přetrvává syntéza pozdních virových proteinů VACV (MULDER *et al.* 1998).

Ačkoliv v buňkách infikovaných VACV faktory eIF4G a eIF4E lokalizují do cytoplasmatických virových továren, protein PABP1 lokalizuje na periferii těchto virových továren a předpokládá se, že se na translaci virových mRNA nepodílí (WALSH *et al.* 2008; XIA *et al.* 2011). Protein PABP1 naopak pravděpodobně hraje roli v inhibici proteosyntézy hostitelské buňky. V časně fázi infekce VACV dochází k produkci krátkých netranslatovaných polyadenylovaných RNA (*polyadenylated RNA sequences*, POLADS), jež inhibují syntézu proteinů

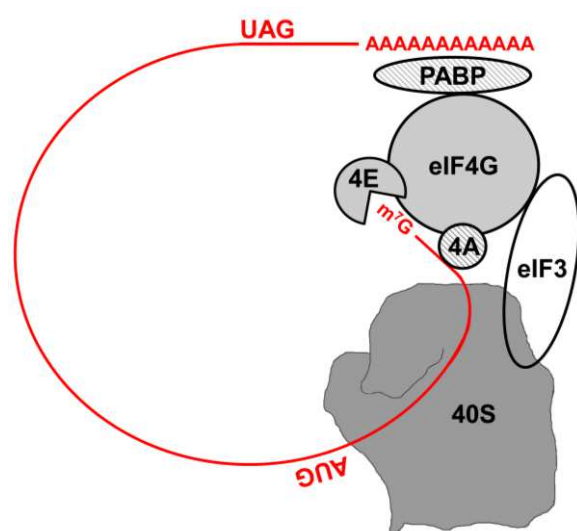
hostitelské buňky svými 3' poly(A) konci, které pravděpodobně sekvestrují PABP1 faktor (BABLANIAN *et al.* 1986; BABLANIAN *et al.* 1987; SU a BABLANIAN 1990; LU a BABLANIAN 1996). POLADS mohou být virového i buněčného původu a předpokládá se, že za syntézu jejich 3' poly(A) konce je zodpovědná sekvenčně nespecifická enzymatická aktivita virové poly(A) polymerázy uvolněné z jádra částice VACV do cytoplasmy buňky během replikace viru (LU a BABLANIAN 1996). POLADS selektivně inhibují translaci pouze buněčných a ne virových transkriptů *in vitro*, což může souviset s tím, že translace virových mRNA může být iniciována *in vitro* i za limitujícího množství PABP1 (BABLANIAN a BANERJEE 1986; BABLANIAN *et al.* 1991).

Současný pohled tedy předpokládá, že translace mRNA VACV je iniciována způsobem závislým na čepičkové struktuře mRNA, ačkoliv translace virových mRNA vykazuje oproti translaci buněčných mRNA sníženou závislost na některých translačních iniciačních faktorech *in vitro* a *in vivo*. To je v přímém rozporu s mými výsledky 5'RACE-PCR analýzy vybraných genů ze všech časových tříd genů VACV, která naznačuje, že výskyt čepičkové struktury je především u transkriptů středních a pozdních genů VACV sporadický. Za předpokladu, že většina z analyzovaných transkriptů může být také translatována, navrhuji níže popsany model.

Předpokládám, že především časně mRNA VACV jsou v buňce s dosud nemodifikovaným translačním aparátem translatovány způsobem závislým na čepičkové struktuře mRNA, kterou také obsahují. V souladu s tímto tvrzením jsou studie dokládající esencialitu eIF4F faktoru pro replikaci VACV (ALDABE *et al.* 1995; WELNOWSKA *et al.* 2009). U transkriptů VACV s 5' poly(A) leaderem je u každé následné třídy genů menší výskyt 5' čepičkové struktury a proto předpokládám, že mechanismy translace virových mRNA se v buňce budou postupně měnit z iniciace závislé na čepičkové struktuře na iniciaci nezávislou na čepičkové struktuře, přičemž tyto změny budou provázeny popsány modifikacemi hostitelského translačního aparátu, které postupně znemožní iniciaci translace závislou na čepičkové struktuře u buněčných mRNA. V souladu s tímto tvrzením jsou například studie dokládající perzistenci produkce pozdních proteinů VACV v buňkách s rozštěpeným eIF4G a postradatelnost eIF4F faktoru pro tvorbu 48S iniciačních ribozomálních komplexů *in vitro* na AUG kodónu 5' polyadenylovaných RNA postrádajících čepičku (MULDER *et al.* 1998; SHIROKIKH a SPIRIN 2008). V souladu s předpokladem iniciace translace nezávislé na čepičkové struktuře u mRNA s 5' poly(A) leadem jsou i výsledky mého školitele dokládající neesencialitu faktoru vázajícího čepičku eIF4E pro translaci mRNA kódujících toxin pGKL plasmidů *in vivo*. Během psaní této dizertační práce a krátce před publikováním našich výsledků se objevila studie, která doložila neesencialitu 5' mRNA čepičky a faktoru eIF4E pro translaci mRNA s krátkým (4–12 nt) 5' poly(A) leadem *in vivo* v buňkách infikovaných VACV (DHUNGEL *et al.* 2017), což sice potvrzuje moji hypotézu o iniciaci translace nezávislé na čepičkové struktuře

u postreplikativních mRNA VACV, ale ke změně pohledu na translaci mRNA VACV tak došlo i bez našeho přičinění. Protože se jedná o zásadní změnu přibližně 40 let starého paradigmatu, autoři svoje výsledky diskutují velmi diplomaticky. Nabízí například variantu, že analogy 5' čepičky, které nemohou být vázány faktorem eIF4E, přítomné na transfekovaných mRNA s 5' poly(A) leadem, mohly být v buňkách odštepeny virovým proteinem D9 či D10 a nahrazeny 5' čepičkou syntetizovanou virovým mRNA *capping* enzymem, přičemž takovéto mRNA by mohly být translatovány způsobem závislým na čepičce (DHUNGEL *et al.* 2017).

Další práce pak ukázala, jak je fosforylaci ribozomálního proteinu RACK1 dosaženo preferenční translace mRNA s dlouhým (25 nt) 5' poly(A) leadem (JHA *et al.* 2017), přičemž v buňkách infikovaných VACV by podle mého názoru mohl být potlačen předpokládaný negativní vliv PABP1 proteinu na translaci mRNA s 5' poly(A) leadem obsahujícím 12 a více následných adenosinů před AUG kodónem díky sekvenci POLADS. Domnívám se také, že společný předek poxvirů a lineárních plasmidů mohl kódovat poly(A) polymerázu, ale díky negativnímu efektu POLADS na genovou expresi hostitele mohlo dojít ke ztrátě příslušných genů v rámci domestikace lineárních plasmidů buňkou. Předpokládám, že v blízké době dojde k prohloubení znalostí mechanismů iniciace translace postreplikativních mRNA VACV s 5' poly(A) leadem, neboť je toto téma cílem grantu řešeného na Kansas State University (YANG 2018). Dosavadní znalosti translace postreplikativních mRNA VACV shrnuje Obrázek 79.



**Obrázek 79:** Zjednodušené schéma iniciace translace transkriptů postreplikativních genů viru *Vaccinia*. Na transkriptu (červeně) je vyznačena 5' čepičková struktura (m<sup>7</sup>G), iniciační kodón (AUG), terminační kodón (UAG) a 3' poly(A) konec (AAAAAAAAAAAA). Malá podjednotka ribozomu (40S) je zobrazena tmavě šedě bez některých dalších translačních iniciačních faktorů, podílejících se na formaci 48S preiniciačního komplexu. Translace postreplikativních mRNA VACV *in vitro* vykazuje sníženou závislost na faktorech eIF4A (4A) a PABP1 (PABP) zobrazených šrafovaně (BABLANIAN *et al.* 1991). Translace postreplikativních mRNA VACV *in vivo* vykazuje sníženou závislost na faktorech eIF4G (eIF4G) a eIF4E (4E) zobrazených světle šedě (MULDER *et al.* 1998; DHUNGEL *et al.* 2017). Výsledky této disertační práce naznačují, že většina mRNA postreplikativních genů neobsahuje 5' čepičkovou strukturu *in vivo*.

Moje hypotéza a výsledky ze systému pGKL plasmidů také vysvětlují, proč všechny postreplikativní promotory VACV obsahují iniciátorový region a proč mutace v INR



postreplikativních promotorů VACV vedou k úplné inhibici exprese reportérového genu. Mutace v INR regionu vedoucí k absenci 5' poly(A) leaderu virové mRNA by i v případě zvýšení četnosti čepičkové struktury na těchto mRNA vedla ke snížení translace těchto mRNA v buňkách v pozdní fázi infekce VACV, ve kterých je narušena translace závislá na čepičkové struktuře. Inhibiční vliv mutací INR regionu postreplikativních promotorů na genovou expresi je však odjakživa interpretován pouze snížením promotorové síly (DAVISON a MOSS 1989b; BALDICK *et al.* 1992). Naproti tomu, mutace INR regionu promotoru genu *K1ORF4* kódujícího podjednotku toxinu pGKL plasmidů vedla jen k částečnému poklesu produkce toxinu, přičemž obdobná mutace v INR regionu promotoru *K1ORF2* prokazatelně vedla k funkční expresi genu *G418<sup>R</sup>*, neboť rezistence ke G418 byla použita pro selekci klonů po transformaci. Předpokládaná změna mechanismů iniciace translace plasmidových mRNA z nezávislosti na závislosti na čepičkové struktuře tak nevede k úplné inhibici genové exprese jako u VACV, neboť lineární plasmidy v hostitelských buňkách s největší pravděpodobností neinhibují translaci závislou na čepičkové struktuře mRNA. Nabízí se však otázka, zda pGKL plasmidy v buňkách neovlivňují lokalizaci translačních iniciačních faktorů. Moje výsledky hmotnostní spektrometrie, naznačující asociaci faktoru eIF4G1 a PABP1 s transkripčním aparátem pGKL plasmidů, připouštějí variantu možného nabohacení translačních iniciačních faktorů v místech transkripce pGKL plasmidů.

### 6.2.3 Složení 3' konců mRNA a terminace transkripce postreplikativních genů viru *Vaccinia*

Tvorba 3' konců transkriptů postreplikativních genů VACV dosud nebyla dopodrobna prozkoumána, hlavně díky výše zmíněné heterogenitě délky 3' konců jednotlivých transkriptů daného genu, která při charakterizaci transkriptů působí značné experimentální potíže. Díky celogenomovým analýzám 5' a 3' konců postreplikativních transkriptů VACV a detekované všudypřítomné tvorbě 3' konců mRNA se předpokládá, že neexistuje specifický sekvenční motiv určující místa tvorby 3' konců mRNA či terminace transkripce postreplikativních genů VACV (YANG *et al.* 2012). Zatím pouze u třech postreplikativních genů VACV byly detekovány definované 3' konce transkriptů (PATEL a PICKUP 1989; AMEGADZIE *et al.* 1992b; D'COSTA *et al.* 2004). U poxvirových genů *A25L* (střední) a *F17R* (pozdní) jsou 3' konce jejich transkriptů generovány endonukleolytickým štěpením, za které je pravděpodobně zodpovědný multifunkční virový protein H5, který je zároveň i pozdním transkripčním faktorem (ANTCZAK *et al.* 1992; D'COSTA *et al.* 2004; D'COSTA *et al.* 2008). K endonukleolytickému štěpení transkriptů dochází v buňkách v pozdních fázích infekce poxvirem a ke štěpení nechochází za podmínek inhibice

replikace virové DNA (ANTCZAK *et al.* 1992; D'COSTA *et al.* 2004). Jelikož promotor středního genu *A25L* vykazuje velmi vysokou aktivitu jako pozdní promotor (YANG *et al.* 2013), je možné, že k endonukleolytickému štěpení dochází především či pouze u pozdních transkriptů viru. Protože transkripty genu *A25L* patří mezi vysoce abundantní virové transkripty v buňce, spekuluje se, že by endonukleolytické štěpení 3' konců mRNA mohlo posilovat expresi vysoce aktivních virových genů (BROYLES 2003).

U transkriptu *A25L* byla identifikována specifická *cis* regulační sekvence (nazvaná AX element) obsahující dvě předpokládané RNA vlásenky, nezbytná pro štěpení 3' konce RNA (HOWARD *et al.* 1999). U transkriptu *F17R* nebyl identifikován element, který by byl AX elementu sekvenčně nebo strukturně podobný (D'COSTA *et al.* 2004). Nové biochemické analýzy virového proteinu H5 *in vitro* ukázaly, že endonukleolytické štěpení RNA vyžadující fosforylovaný protein H5 poskytuje volný 3' OH konec, který by mohl být využit jako templát virovou poly(A) polymerázou pro syntézu 3' poly(A) konce mRNA (KAY *et al.* 2013). Protein H5 preferenčně váže dvouvláknové nukleové kyseliny způsobem nezávislým na jejich sekvenci *in vitro* (KAY *et al.* 2013), což by mohlo vysvětlit, proč nebyla u transkriptu *F17R* identifikována sekvence či struktura podobná AX elementu, neboť jako substrát pro štěpení H5 proteinem by teoreticky mohl postačit stonek jakékoliv RNA vlásenky. Pro pozdní transkripty genů VACV, jejichž 3' konce jsou generovány endonukleolytickým štěpením, byla navržena terminace transkripce modelem torpéda, kdy kotranskripční štěpení nascentní virové mRNA proteinem H5 způsobí destabilizaci elongačního komplexu RNAP (KAY *et al.* 2013). Kromě dobře charakterizovaných příkladů dvou poxvirových genů však není jasné, kolika postreplikativních genů VACV se štěpení 3' konce RNA týká.

Tvorba 3' konců mRNA vybraných postreplikativních genů poxvirů analogicky připomíná situaci u charakterizovaných 3' konců mRNA fylogeneticky příbuzných virů čeledi *Mimiviridae*, kde se na 3' koncích transkriptů většiny (82 %) genů všech časových tříd podle celogenomové bioinformatické analýzy nachází RNA vlásenky (BYRNE *et al.* 2009). V rámci těchto RNA vlásenek pak zřejmě dochází k endonukleolytickému štěpení virovým proteinem s předpokládanou strukturou podobnou ribonukleáze III, přičemž štěpení poskytne volný 3' OH konec pro syntézu 3' poly(A) konce mRNA virovou poly(A) polymerázou (PRIET *et al.* 2015). Díky celogenomové predikci RNA vlásenek na 3' koncích mRNA většiny genů a identifikaci předpokládaných orthologů mimivirového proteinu podobnému ribonukleáze III je nyní model tvorby 3' konců mRNA endonukleolytickým štěpením předpokládán i pro další nukleocytoplasmatické viry čeledí *Megaviridae*, *Pithoviridae* a *Marseilleviridae* (ARSLAN *et al.* 2011; LEGENDRE *et al.* 2014; FABRE *et al.* 2017). V neposlední řadě AU bohaté vlásenky na 3' koncích RNA byly detekovány u některých transkriptů virů čeledí *Ascoviridae* a *Iridoviridae*, kde se zřejmě

podílí na tvorbě 3' konců mRNA buď endonukleolytickým štěpením nebo indukci terminace transkripce (SALEM *et al.* 2008; ÍNCE *et al.* 2017; ZAGHLOUL *et al.* 2017).

Z výše uvedených recentních prací vyplývá, že RNA vlásenky zřejmě běžně hrají roli při formaci 3' konců mRNA u nukleocytoplasmatických virů většiny čeledí. U virů s velkými genomy z čeledí *Mimiviridae*, *Megaviridae*, *Pithoviridae* a *Marseilleviridae* mají RNA vlásenky poměrně uniformní délku a nachází se většinou v mezigenových oblastech, díky čemuž se poměrně dobře predikují *in silico* (BYRNE *et al.* 2009; ARSLAN *et al.* 2011; LEGENDRE *et al.* 2014; FABRE *et al.* 2017). U virů čeledi *Iridoviridae* mají AU bohaté RNA vlásenky rozdílnou délku, ale přibližně polovina z nich má konzervovaný komplementární pentanukleotidový motiv tvořící bázi stonku vlásenky (ÍNCE *et al.* 2017). Výsledky této disertační práce naznačují, že metodické obtíže při charakterizaci 3' konců mRNA postreplikativních genů VACV působené heterogenitou délky jednotlivých transkriptů lze běžně vyřešit použitím 3' RACE-PCR analýzy. Touto metodikou byly zmapovány sekvence 3' konců mRNA čtyř vybraných postreplikativních genů VACV, které se nachází poblíž předpokládaných slabých RNA vlásenek s mediánem Gibbsovy volné energie  $-7,9$  kcal/mol. Je tak možné, že tyto RNA vlásenky budou mít roli při formaci 3' konců mRNA postreplikativních genů VACV.

Regulační AX element modelového transkriptu genu *A25L*, obsahující dvě předpokládané RNA vlásenky, má srovnatelnou Gibbsovu volnou energii ( $-8,5$  kcal/mol) jako předpokládané RNA vlásenky identifikované touto prací. Je tak možné, že detekované 3' konce mRNA vybraných postreplikativních genů VACV vznikly endonukleolytickým štěpením. Osobně si ale myslím, že RNA vlásenky mohou mít při tvorbě 3' konců mRNA postreplikativních genů VACV dvojí roli. Předpokládám, že slabé RNA vlásenky obecně dokážou indukovat terminaci transkripce postreplikativních genů VACV a ve fázi exprese pozdních genů VACV mohou být některé tyto RNA vlásenky štěpeny na definovaném místě virovým proteinem H5. Moji hypotézu podporuje několik následujících zjištění.

Slabé RNA vlásenky dokážou ovlivnit účinnost i umístění terminace transkripce časných genů VACV *in vitro*, přičemž se předpokládá, že slabé RNA vlásenky by mohly indukovat terminaci transkripce časných genů VACV, u kterých chybí kanonický 5'-UUUUUNU-3' terminační motiv (TATE a GOLLNICK 2015). Touto prací identifikované RNA vlásenky transkriptu genu *A27L* VACV dokázaly indukovat formování 3' konců mRNA na obdobné pozici po umístění do pGKL plasmidu *in vivo*, přičemž pGKL plasmidy pravděpodobně nedisponují proteinem homologním H5 proteinu VACV. Ačkoliv ke štěpení modelových transkriptů genů *A25L* a *F17R* poxvirů dochází predominantně na jednom konkrétním nukleotidu *in vitro* (ANTCZAK *et al.* 1992; D'COSTA *et al.* 2004), celogenomovou transkriptomickou analýzou VACV byly identifikovány terminální nukleotidy nachazející se

v okolí štěpících míst těchto transkriptů ve stejné transkripční orientaci (YANG *et al.* 2012), přičemž tyto detekované 3' konce mRNA mohly vzniknout spíše terminací transkripce než endonukleolytickým štěpením, neboť i v případě popsaného štěpícího místa transkriptu *F17R* se mi poblíž podařilo identifikovat předpokládanou slabou RNA vlásenkovou strukturu (data nezobrazena). Budoucí studie musí objasnit hypotetický vliv slabých RNA vlásenek na tvorbu 3' konců transkriptů postreplikativních genů u poxvirů.

## 7 SOUHRN

Hlavním cílem této práce bylo charakterizovat transkripční aparát lineárních plasmidů kvasinky *Kluyveromyces lactis* bioinformatickými, genetickými a biochemickými metodami.

Bioinformatická charakterizace *in silico*:

- Nekanonická RNA polymeráza a plasmidové promotory mají původ shodný s poxviry.
- Iniciátorový region promotoru (TAAAN) je zodpovědný za netemplátovou 5' polyadenylaci některých plasmidových transkriptů.
- Struktury slabých RNA vlásenek a zřejmě i 3' oligo(U) konec RNA se podílí na definované tvorbě 3' konců plasmidových transkriptů.
- Otevřený čtecí rámec 9 plasmidu pGKL2 kóduje předpokládaný *scaffold* protein homologní k proteinu D13 viru *Vaccinia*.

Genetická a biochemická charakterizace:

- Methyltransferázová doména mRNA *capping* enzymu je esenciální pro fungování plasmidů *in vivo*.
- Plasmidové transkripty obsahují 5' čepičkovou strukturu *in vivo*.
- Všechny složky transkripčního aparátu plasmidů interagují *in vivo*; s transkripčním komplexem interaguje předpokládaný *scaffold* protein *in vivo*.
- RNA polymerázovou aktivitu purifikovaného transkripčního komplexu se nepodařilo detekovat *in vitro*.
- Předpokládaná helikáza specificky interaguje s plasmidovou DNA *in vivo*.
- Podařilo se prokázat přímý vliv iniciátorového regionu promotoru na 5' polyadenylaci plasmidových transkriptů *in vivo*.
- Podařilo se prokázat přímý vliv RNA vlásenky na tvorbu 3' konců plasmidových transkriptů *in vivo*.

Dalšími cíli této práce, v návaznosti na určení předpokládaného původu transkripčního aparátu lineárních plasmidů, byla charakterizace nepřekládaných oblastí transkriptů viru *Vaccinia* z čeledi *Poxviridae* bioinformatickými, genetickými a biochemickými metodami.

Bioinformatická, genetická a biochemická charakterizace:

- Frekvence výskytu 5' čepičky je odlišná u poxvirových transkriptů různých časových tříd genů.
- 3' konce postreplikativních transkriptů poxvirů vznikajících *in vivo* lze detekovat.
- Na tvorbě 3' konců postreplikativních transkriptů poxvirů *in vivo* se zřejmě podílí struktury slabých RNA vlásenek.

## 8 SEZNAM PUBLIKACÍ

### 8.1 PUBLIKACE

**Sýkora, M.**, Pospíšek, M., Novák, J., Mrvová, S., Krásný, L., Vopálenský, V. (2018): Transcription apparatus of the yeast virus-like elements: Architecture, function, and evolutionary origin. *PLoS Pathogens*, 14(10), e1007377.

Vopálenský, V., **Sýkora, M.**, Mašek, T., Pospíšek, M. (2019): Messenger RNAs of the yeast virus-like elements contain non-templated 5' poly(A) leaders and their expression is independent of eIF4E and Pab1. *Frontiers in Microbiology*, v revizích.

Cackett, G., Sheppard, C., Portugal, R., Malecki, M., **Sýkora, M.**, Fouqueau, T., Blombach, F., Matelska, D., Dixon, L., Bähler, J., Werner, F. (připravovaný rukopis): Transcription Start Site Mapping, Promoter Analysis and Transcriptome of African Swine Fever Virus.

**Sýkora, M.**, Pospíšek, M., Mělková, Z., Vopálenský, V. (plánovaný rukopis): 3' mRNA ends of Vaccinia virus postreplicative transcripts are located close to putative RNA stem loop structures.

### 8.2 KONFERENCE

#### 8.2.1 Přednášky

**Sýkora, M.** Initial characterization of a putative non-canonical RNA polymerase encoded by yeast cytoplasmic linear plasmids. miniRNA meeting, Želiv, Česká republika (listopad 2013).

**Sýkora, M.**, Pospíšek, M., Vopálenský, V. Transcription apparatus of the yeast linear plasmids points to its poxviral origin. International RNA Club meeting, Praha, Česká republika (září 2014).

**Sýkora, M.** The relationship between 5' poly(A) leader and 5' cap in mRNAs of yeast linear plasmids and poxviruses. miniRNA meeting, Želiv, Česká republika (listopad 2015).

#### 8.2.2 Postery

Vopálenský, V., **Sýkora, M.**, Pospíšek, M. Unique world of linear plasmids of the yeast *Kluyveromyces lactis*. 27. Kongres Československé společnosti mikrobiologické, Praha, Česká republika (září 2016).



## 8.3 PODÍL STUDENTA NA PUBLIKACÍCH

### Publikace I:

**Transcription apparatus of the yeast virus-like elements: Architecture, function, and evolutionary origin.**

Michal Sýkora, Martin Pospíšek, Josef Novák, Silvia Mrvová, Libor Krásný a Václav Vopálenský

Podíl studenta: 85 %; Navrhnul jsem a naplánoval všechny experimenty, většinu z nich jsem provedl sám. Zanalyzoval jsem výsledky a napsal rukopis. Řídil jsem závěrečné úpravy a publikační proces rukopisu.

### Publikace II:

**Messenger RNAs of the yeast virus-like elements contain non-templated 5' poly(A) leaders and their expression is independent of eIF4E and Pab1.**

Václav Vopálenský, Michal Sýkora, Tomáš Mašek a Martin Pospíšek

Podíl studenta: 25 %; Podílel jsem se na návrhu a provedení experimentů v Obrázku 5, 7, S1 a S3. Provedl jsem analýzu dat vedoucí k Obrázku 3, 4, 6 a S2. Též jsem se podílel na analýze všech výsledků a přispěl jsem k sepisování a závěrečným úpravám textu rukopisu.

### Publikace III:

**Transcription Start Site Mapping, Promoter Analysis and Transcriptome of African Swine Fever Virus.**

Gweny Cackett, Carol Sheppard, Raquel Portugal, Michal Malecki, Michal Sýkora, Thomas Fouqueau, Fabian Blombach, Dorota Matelska, Linda Dixon, Jürg Bähler, Finn Werner

Podíl studenta: 5 %; Při bioinformatické analýze nukleotidů obklopujících experimentálně zjištěná majoritní místa transkripčních startů diferencially exprimovaných genů viru afrického moru prasat (*Asfarviridae*) jsem detekoval předpokládaný promotorový element časných genů tohoto viru, jenž je svým umístěním a sekvencí podobný UCS elementu pGKL plasmidů a UCE elementu VACV. Rukopis je připravován.

### Publikace IV:

**3' mRNA ends of Vaccinia virus postreplicative transcripts are located close to putative RNA stem loop structures.**

Michal Sýkora, Martin Pospíšek, Zora Mělková a Václav Vopálenský

Podíl studenta: 80 %; Navrhnul jsem a naplánoval všechny experimenty. Přímoou práci s virem *Vaccinia* provedla Zora Mělková. Následné experimenty využívající virovou RNA jsem provedl sám. Zanalyzoval jsem výsledky. Příprava rukopisu je plánována.

Tímto potvrzuji, že autor práce, Michal Sýkora, významně přispěl k výše uvedeným publikacím. V případě jeho prvoautorské publikace (nesdílené) vykonal hlavní část experimentální práce a přispěl k přípravě rukopisu.

Mgr. Václav Vopálenský, Ph.D.

## 9 SEZNAM POUŽITÉ LITERATURY

- Abu-Daya, A., Brown, P. M., Fox, K. R. (1995): DNA sequence preferences of several AT-selective minor groove binding ligands. *Nucleic Acids Research*, 23(17), 3385-3392.
- Ahn, B. Y., Gershon, P. D., Jones, E. V., Moss, B. (1990a): Identification of rpo30, a Vaccinia Virus RNA Polymerase Gene with Structural Similarity to a Eucaryotic Transcription Elongation Factor. *Molecular and Cellular Biology*, 10(10), 5433-5441.
- Ahn, B. Y., Gershon, P. D., Moss, B. (1994): RNA polymerase-associated protein Rap94 confers promoter specificity for initiating transcription of vaccinia virus early stage genes. *Journal of Biological Chemistry*, 269(10), 7552-7557.
- Ahn, B. Y., Jones, E. V., Moss, B. (1990b): Identification of the Vaccinia Virus Gene Encoding an 18-Kilodalton Subunit of RNA Polymerase and Demonstration of a 5' Poly(A) Leader on Its Early Transcript. *Journal of Virology*, 64(6), 3019-3024.
- Ahn, B. Y., Moss, B. (1989): Capped Poly(A) Leaders of Variable Lengths at the 5' Ends of Vaccinia Virus Late mRNAs. *Journal of Virology*, 63(1), 226-232.
- Ahn, B. Y., Moss, B. (1992): RNA polymerase-associated transcription specificity factor encoded by vaccinia virus. *Proceedings of the National Academy of Sciences of the United States of America*, 89(8), 3536-3540.
- Ahn, B. Y., Rosel, J., Cole, N. B., Moss, B. (1992): Identification and Expression of rpo19, a Vaccinia Virus Gene Encoding a 19-Kilodalton DNA-Dependent RNA Polymerase Subunit. *Journal of Virology*, 66(2), 971-982.
- Aldabe, R., Feduchi, E., Novoa, I., Carrasco, L. (1995): Efficient cleavage of p220 by poliovirus 2Apro expression in mammalian cells: effects on vaccinia virus. *Biochemical and Biophysical Research Communications*, 215(3), 928-936.
- Altmann, M., Handschin, C., Trachsel, H. (1987): mRNA Cap-Binding Protein: Cloning of the Gene Encoding Protein Synthesis Initiation Factor eIF-4E from *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 7(3), 998-1003.
- Amegadzie, B. Y., Ahn, B. Y., Moss, B. (1991): Identification, Sequence, and Expression of the Gene Encoding a Mr 35,000 Subunit of the Vaccinia Virus DNA-dependent RNA Polymerase. *Journal of Biological Chemistry*, 266(21), 13712-13718.
- Amegadzie, B. Y., Ahn, B. Y., Moss, B. (1992a): Characterization of a 7-Kilodalton Subunit of Vaccinia Virus DNA-Dependent RNA Polymerase with Structural Similarities to the Smallest Subunit of Eukaryotic RNA Polymerase II. *Journal of Virology*, 66(5), 3003-3010.
- Amegadzie, B. Y., Sisler, J. R., Moss, B. (1992b): Frame-Shift Mutations within the Vaccinia Virus A-Type Inclusion Protein Gene. *Virology*, 186(2), 777-782.
- Ansarah-Sobrinho, C., Moss, B. (2004): Role of the I7 Protein in Proteolytic Processing of Vaccinia Virus Membrane and Core Components. *Journal of Virology*, 78(12), 6335-6343.
- Antczak, J. B., Patel, D. D., Ray, C. A., Ink, B. S., Pickup, D. J. (1992): Site-specific RNA cleavage generates the 3' end of a poxvirus late mRNA. *Proceedings of the National Academy of Sciences of the United States of America*, 89(24), 12033-12037.
- Arslan, D., Legendre, M., Seltzer, V., Abergel, C., Claverie, J. M. (2011): Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proceedings of the National Academy of Sciences of the United States of America*, 108(42), 17486-17491.
- Bablanian, R., Banerjee, A. K. (1986): Poly(riboadenylic acid) preferentially inhibits in vitro translation of cellular mRNAs compared with vaccinia virus mRNAs: Possible role in vaccinia virus cytopathology. *Proceedings of the National Academy of Sciences*, 83(5), 1290-1294.
- Bablanian, R., Coppola, G., Masters, P. S., Banerjee, A. K. (1986): Characterization of Vaccinia Virus Transcripts Involved in Selective Inhibition of Host Protein Synthesis. *Virology*, 148(2), 375-380.
- Bablanian, R., Goswami, S. K., Esteban, M., Banerjee, A. K. (1987): Selective Inhibition of Protein Synthesis by Synthetic and Vaccinia Virus-Core Synthesized Poly(riboadenylic acids). *Virology*, 161(2), 366-373.
- Bablanian, R., Goswami, S. K., Esteban, M., Banerjee, A. K., Merrick, W. C. (1991): Mechanism of Selective Translation of Vaccinia Virus mRNAs: Differential Role of Poly(A) and Initiation Factors in the Translation of Viral and Cellular mRNAs. *Journal of Virology*, 65(8), 4449-4460.
- Bag, J., Wu, J. (1996): Translational Control of Poly(A)-Binding Protein Expression. *European Journal of Biochemistry*, 237(1), 143-152.
- Bahar, Mohammad W., Graham, Stephen C., Stuart, David I., Grimes, Jonathan M. (2011): Insights into the Evolution of a Complex Virus from the Crystal Structure of Vaccinia Virus D13. *Structure*, 19(7-12), 1011-1020.
- Bajszár, G., Wittek, R., Weir, J. P., Moss, B. (1983): Vaccinia Virus Thymidine Kinase and Neighboring Genes: mRNAs and Polypeptides of Wild-Type Virus and Putative Nonsense Mutants. *Journal of Virology*, 45(1), 62-72.

- Baldick, C. J., Jr., Cassetti, M. C., Harris, N., Moss, B. (1994): Ordered Assembly of a Functional Preinitiation Transcription Complex, Containing Vaccinia Virus Early Transcription Factor and RNA Polymerase, on an Immobilized Template. *Journal of Virology*, 68(9), 6052-6056.
- Baldick, C. J., Jr., Keck, J. G., Moss, B. (1992): Mutational Analysis of the Core, Spacer, and Initiator Regions of Vaccinia Virus Intermediate-Class Promoters. *Journal of Virology*, 66(8), 4710-4719.
- Baldick, C. J., Jr., Moss, B. (1993): Characterization and Temporal Regulation of mRNAs Encoded by Vaccinia Virus Intermediate-Stage Genes. *Journal of Virology*, 67(6), 3515-3527.
- Barbosa, E., Moss, B. (1978): mRNA(nucleoside-2')-methyltransferase from Vaccinia Virus. Characteristics and substrate specificity. *Journal of Biological Chemistry*, 253(21), 7698-7702.
- Baroudy, B. M., Moss, B. (1980): Purification and Characterization of a DNA-dependent RNA Polymerase from Vaccinia Virions. *Journal of Biological Chemistry*, 255(9), 4372-4380.
- Baroudy, B. M., Moss, B. (1982): Sequence homologies of diverse length tandem repetitions near ends of vaccinia virus genome suggest unequal crossing over. *Nucleic Acids Research*, 10(18), 5673-5679.
- Baroudy, B. M., Venkatesan, S., Moss, B. (1982): Incompletely Base-Paired Flip-Flop Terminal Loops Link the Two DNA Strands of the Vaccinia Virus Genome into One Uninterrupted Polynucleotide Chain. *Cell*, 28(2), 315-324.
- Baxby, D. (1999): The origins of vaccinia virus - An even shorter rejoinder. *Social History of Medicine*, 12(1), 139-139.
- Baylis, S. A., Twigg, S. R., Vydelingum, S., Dixon, L. K., Smith, G. L. (1993): Three African swine fever virus genes encoding proteins with homology to putative helicases of vaccinia virus. *Journal of General Virology*, 74 (Pt 9), 1969-1974.
- Bayliss, C. D., Condit, R. C. (1993): Temperature-Sensitive Mutants in the Vaccinia Virus A18R Gene Increase Double-Stranded RNA Synthesis as a Result of Aberrant Viral Transcription. *Virology*, 194(1), 254-262.
- Bayliss, C. D., Condit, R. C. (1995): The Vaccinia Virus A18R Gene Product Is a DNA-dependent ATPase. *Journal of Biological Chemistry*, 270(4), 1550-1556.
- Bayliss, C. D., Smith, G. L. (1996): Vaccinia Virion Protein I8R Has both DNA and RNA Helicase Activities: Implications for Vaccinia Virus Transcription. *Journal of Virology*, 70(2), 794-800.
- Beaud, G., Beaud, R. (2000): Temperature-Dependent Phosphorylation State of the H5R Protein Synthesised at the Early Stage of Infection in Cells Infected with Vaccinia Virus ts Mutants of the B1R and F10L Protein Kinases. *Intervirology*, 43(1), 67-70.
- Beaud, G., Beaud, R., Leader, D. P. (1995): Vaccinia Virus Gene H5R Encodes a Protein That Is Phosphorylated by the Multisubstrate Vaccinia Virus B1R Protein Kinase. *Journal of Virology*, 69(3), 1819-1826.
- Bellaousov, S., Reuter, J. S., Seetin, M. G., Mathews, D. H. (2013): RNAstructure: web servers for RNA secondary structure prediction and analysis. *Nucleic Acids Research*, 41(Web Server issue), W471-W474.
- Benarroch, D., Smith, P., Shuman, S. (2008): Characterization of a trifunctional mimivirus mRNA capping enzyme and crystal structure of the RNA triphosphatase domain. *Structure*, 16(4), 501-512.
- Berset, C., Zurbriggen, A., Djafarzadeh, S., Altmann, M., Trachsel, H. (2003): RNA-binding activity of translation initiation factor eIF4G1 from *Saccharomyces cerevisiae*. *RNA*, 9(7), 871-880.
- Bertholet, C., Stocco, P., Van Meir, E., Wittek, R. (1986): Functional analysis of the 5' flanking sequence of a vaccinia virus late gene. *The EMBO Journal*, 5(8), 1951-1957.
- Bertholet, C., Van Meir, E., ten Heggeler-Bordier, B., Wittek, R. (1987): Vaccinia Virus Produces Late mRNAs by Discontinuous Synthesis. *Cell*, 50(2), 153-162.
- Bisht, H., Weisberg, A. S., Szajner, P., Moss, B. (2009): Assembly and Disassembly of the Capsid-Like External Scaffold of Immature Virions during Vaccinia Virus Morphogenesis. *Journal of Virology*, 83(18), 9140-9150.
- Black, E. P., Condit, R. C. (1996): Phenotypic Characterization of Mutants in Vaccinia Virus Gene G2R, a Putative Transcription Elongation Factor. *Journal of Virology*, 70(1), 47-54.
- Black, E. P., Moussatche, N., Condit, R. C. (1998): Characterization of the Interactions among Vaccinia Virus Transcription Factors G2R, A18R, and H5R. *Virology*, 245(2), 313-322.
- Blaisonneau, J., Nosek, J., Fukuhara, H. (1999): Linear DNA plasmid pPK2 of *Pichia kluyveri*: Distinction Between Cytoplasmic and Mitochondrial Linear Plasmids in Yeasts. *Yeast*, 15(9), 781-791.
- Boone, R. F., Moss, B. (1977): Methylated 5'-Terminal Sequences of Vaccinia Virus mRNA Species Made in Vivo at Early and Late Times after Infection. *Virology*, 79(1), 67-80.
- Boone, R. F., Moss, B. (1978): Sequence Complexity and Relative Abundance of Vaccinia Virus mRNA's Synthesized In Vivo and In Vitro. *Journal of Virology*, 26(3), 554-569.
- Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., Robert, C., Azza, S., Sun, S., Rossmann, M. G., Suzan-Monti, M., La Scola, B., Koonin, E. V., Raoult, D. (2009): Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proceedings of the National Academy of Sciences of the United States of America*, 106(51), 21848-21853.
- Boyle, K. A., Greseth, M. D., Traktman, P. (2015): Genetic Confirmation that the H5 Protein Is Required for Vaccinia Virus DNA Replication. *Journal of Virology*, 89(12), 6312-6327.
- Brown, N. G., Nick Morrice, D., Beaud, G., Hardie, G., Leader, D. P. (2000): Identification of sites phosphorylated by the vaccinia virus B1R kinase in viral protein H5R. *BMC Biochemistry*, 1, 2.

- Broyles, S. S. (1991): A Role for ATP Hydrolysis in Vaccinia Virus Early Gene Transcription - Dissociation of the Early Transcription Factor-Promoter Complex. *Journal of Biological Chemistry*, 266(23), 15545-15548.
- Broyles, S. S. (2003): Vaccinia virus transcription. *Journal of General Virology*, 84(Pt 9), 2293-2303.
- Broyles, S. S., Fesler, B. S. (1990): Vaccinia Virus Gene Encoding a Component of the Viral Early Transcription Factor. *Journal of Virology*, 64(4), 1523-1529.
- Broyles, S. S., Knutson, B. A. (2010): Poxvirus transcription. *Future Virology*, 5(5), 639-650.
- Broyles, S. S., Kremer, M., Knutson, B. A. (2004): Antiviral Activity of Distamycin A against Vaccinia Virus Is the Result of Inhibition of Postreplicative mRNA Synthesis. *Journal of Virology*, 78(4), 2137-2141.
- Broyles, S. S., Li, J. (1993): The Small Subunit of the Vaccinia Virus Early Transcription Factor Contacts the Transcription Promoter DNA. *Journal of Virology*, 67(9), 5677-5680.
- Broyles, S. S., Li, J., Moss, B. (1991): Promoter DNA contacts made by the vaccinia virus early transcription factor. *Journal of Biological Chemistry*, 266(23), 15539-15544.
- Broyles, S. S., Liu, X., Zhu, M., Kremer, M. (1999): Transcription Factor YY1 Is a Vaccinia Virus Late Promoter Activator. *Journal of Biological Chemistry*, 274(50), 35662-35667.
- Broyles, S. S., Moss, B. (1986): Homology between RNA polymerases of poxviruses, prokaryotes, and eukaryotes: nucleotide sequence and transcriptional analysis of vaccinia virus genes encoding 147-kDa and 22-kDa subunits. *Proceedings of the National Academy of Sciences of the United States of America*, 83(10), 3141-3145.
- Broyles, S. S., Moss, B. (1987a): Identification of the Vaccinia Virus Gene Encoding Nucleoside Triphosphate Phosphohydrolase I, a DNA-Dependent ATPase. *Journal of Virology*, 61(5), 1738-1742.
- Broyles, S. S., Moss, B. (1987b): Sedimentation of an RNA Polymerase Complex from Vaccinia Virus That Specifically Initiates and Terminates Transcription. *Molecular and Cellular Biology*, 7(1), 7-14.
- Broyles, S. S., Moss, B. (1988): DNA-dependent ATPase Activity Associated with Vaccinia Virus Early Transcription Factor. *Journal of Biological Chemistry*, 263(22), 10761-10765.
- Broyles, S. S., Yuen, L., Shuman, S., Moss, B. (1988): Purification of a Factor Required for Transcription of Vaccinia Virus Early Genes. *Journal of Biological Chemistry*, 263(22), 10754-10760.
- Bugert, J. J., Melquiot, N. V., Darai, G. (2000): Mapping of mRNA Transcripts in the Genome of Molluscum Contagiosum Virus: Transcriptional Analysis of the Viral Slam Gene Family. *Virus Genes*, 21(3), 189-192.
- Burd, C. G., Matunis, E. L., Dreyfuss, G. (1991): The Multiple RNA-Binding Domains of the mRNA Poly(A)-Binding Protein Have Different RNA-Binding Activities. *Molecular and Cellular Biology*, 11(7), 3419-3424.
- Byrd, A. K., Raney, K. D. (2012): Superfamily 2 helicases. *Frontiers in Bioscience-Landmark*, 17, 2070-2088.
- Byrd, C. M., Bolken, T. C., Hruby, D. E. (2002): The vaccinia virus I7L gene product is the core protein proteinase. *Journal of Virology*, 76(17), 8973-8976.
- Byrne, D., Grzela, R., Lartigue, A., Audic, S., Chenivresse, S., Encinas, S., Claverie, J. M., Abergel, C. (2009): The polyadenylation site of Mimivirus transcripts obeys a stringent 'hairpin rule'. *Genome Research*, 19(7), 1233-1242.
- Cairns, J. (1960): The Initiation of Vaccinia Infection. *Virology*, 11, 603-623.
- Cao, S., Dhungel, P., Yang, Z. (2017): Going against the Tide: Selective Cellular Protein Synthesis during Virally Induced Host Shutoff. *Journal of Virology*, 91(17), e00071-17.
- Carpenter, M. S., DeLange, A. M. (1992): Identification of a Temperature-Sensitive Mutant of Vaccinia Virus Defective in Late but Not Intermediate Gene Expression. *Virology*, 188(1), 233-244.
- Carter, G. C., Rodger, G., Murphy, B. J., Law, M., Krauss, O., Hollinshead, M., Smith, G. L. (2003): Vaccinia virus cores are transported on microtubules. *Journal of General Virology*, 84(Pt 9), 2443-2458.
- Cassetti, M. C., Moss, B. (1996): Interaction of the 82-kDa subunit of the vaccinia virus early transcription factor heterodimer with the promoter core sequence directs downstream DNA binding of the 70-kDa subunit. *Proceedings of the National Academy of Sciences of the United States of America*, 93(15), 7540-7545.
- Cochran, M. A., Puckett, C., Moss, B. (1985): In Vitro Mutagenesis of the Promoter Region for a Vaccinia Virus Gene: Evidence for Tandem Early and Late Regulatory Signals. *Journal of Virology*, 54(1), 30-37.
- Colby, C. (1971): Mechanism of Synthesis of Vaccinia Virus Double-Stranded Ribonucleic Acid In Vivo and In Vitro. *Journal of Virology*, 7(1), 71-76.
- Colson, P., de Lamballerie, X., Fournous, G., Raoult, D. (2012): Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. *Intervirology*, 55(5), 321-332.
- Colson, P., Pagnier, I., Yoosuf, N., Fournous, G., La Scola, B., Raoult, D. (2013): "Marseilleviridae", a new family of giant viruses infecting amoebae. *Archives of Virology*, 158(4), 915-920.
- Condit, R. C., Easterly, R., Pacha, R. F., Fathi, Z., Meis, R. J. (1991): A Vaccinia Virus Isatin- $\beta$ -Thiosemicarbazone Resistance Mutation Maps in the Viral Gene Encoding the 132-kDa Subunit of RNA Polymerase. *Virology*, 185(2), 857-861.
- Condit, R. C., Lewis, J. I., Quinn, M., Christen, L. M., Niles, E. G. (1996a): Use of Lysolecithin-Permeabilized Infected-Cell Extracts to Investigate the in Vitro Biochemical Phenotypes of Poxvirus ts Mutations Altered in Viral Transcription Activity. *Virology*, 218(1), 169-180.
- Condit, R. C., Niles, E. G. (2002): Regulation of viral transcription elongation and termination during vaccinia virus infection. *Biochimica et Biophysica Acta*, 1577(2), 325-336.
- Condit, R. C., Xiang, Y., Lewis, J. I. (1996b): Mutation of Vaccinia Virus Gene G2R Causes Suppression of Gene A18R ts Mutants: Implications for Control of Transcription. *Virology*, 220(1), 10-19.

- Cong, P., Shuman, S. (1992): Methyltransferase and Subunit Association Domains of Vaccinia Virus mRNA Capping Enzyme. *Journal of Biological Chemistry*, 267(23), 16424-16429.
- Cong, P., Shuman, S. (1993): Covalent Catalysis in Nucleotidyl Transfer. A KTDG motif essential for enzyme-GMP complex formation by mRNA capping enzyme is conserved at the active sites of RNA and DNA ligases. *Journal of Biological Chemistry*, 268(10), 7256-7260.
- Cong, P., Shuman, S. (1995): Mutational Analysis of mRNA Capping Enzyme Identifies Amino Acids Involved in GTP Binding, Enzyme-Guanylate Formation, and GMP Transfer to RNA. *Molecular and Cellular Biology*, 15(11), 6222-6231.
- Cooper, J. A., Wittek, R., Moss, B. (1981): Extension of the Transcriptional and Translational Map of the Left End of the Vaccinia Virus Genome to 21 Kilobase Pairs. *Journal of Virology*, 39(3), 733-745.
- Cormack, B. P., Bertram, G., Egerton, M., Gow, N. A., Falkow, S., Brown, A. J. (1997): Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in *Candida albicans*. *Microbiology*, 143 (Pt 2), 303-311.
- Cox, J., Hein, M. Y., Lubner, C. A., Paron, I., Nagaraj, N., Mann, M. (2014): Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Molecular & Cellular Proteomics*, 13(9), 2513-2526.
- Cox, J., Mann, M. (2012): 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. *BMC Bioinformatics*, 13(Suppl 16), S12-S12.
- Cresawn, S. G., Condit, R. C. (2007): A targeted approach to identification of vaccinia virus postreplicative transcription elongation factors: Genetic evidence for a role of the H5R gene in vaccinia transcription. *Virology*, 363(2), 333-341.
- Cresawn, S. G., Prins, C., Latner, D. R., Condit, R. C. (2007): Mapping and phenotypic analysis of spontaneous isatin- $\beta$ -thiosemicarbazone resistant mutants of vaccinia virus. *Virology*, 363(2), 319-332.
- Crooks, G. E., Hon, G., Chandonia, J. M., Brenner, S. E. (2004): WebLogo: A Sequence Logo Generator. *Genome Research*, 14(6), 1188-1190.
- Cudmore, S., Cossart, P., Griffiths, G., Way, M. (1995): Actin-based motility of vaccinia virus. *Nature*, 378(6557), 636-638.
- Cyrklaff, M., Risco, C., Fernández, J. J., Jiménez, M. V., Estéban, M., Baumeister, W., Carrascosa, J. L. (2005): Cryo-electron tomography of vaccinia virus. *Proceedings of the National Academy of Sciences of the United States of America*, 102(8), 2772-2777.
- d'Aubenton Carafa, Y., Brody, E., Thermes, C. (1990): Prediction of Rho-independent *Escherichia coli* Transcription Terminators. A Statistical Analysis of their RNA Stem-Loop Structures. *Journal of Molecular Biology*, 216(4), 835-858.
- D'Costa, S. M., Antczak, J. B., Pickup, D. J., Condit, R. C. (2004): Post-transcription cleavage generates the 3' end of F17R transcripts in vaccinia virus. *Virology*, 319(1), 1-11.
- D'Costa, S. M., Bainbridge, T. W., Condit, R. C. (2008): Purification and Properties of the Vaccinia Virus mRNA Processing Factor. *Journal of Biological Chemistry*, 283(9), 5267-5275.
- D'Costa, S. M., Bainbridge, T. W., Kato, S. E., Prins, C., Kelley, K., Condit, R. C. (2010): Vaccinia H5 is a multifunctional protein involved in viral DNA replication, postreplicative gene transcription, and virion morphogenesis. *Virology*, 401(1), 49-60.
- Da Silva, M., Upton, C. (2009): Vaccinia Virus G8R Protein: A Structural Ortholog of Proliferating Cell Nuclear Antigen (PCNA). *PLoS One*, 4(5), e5479.
- Dailidienė, D., Tan, S., Ogura, K., Zhang, M., Lee, A. H., Severinov, K., Berg, D. E. (2007): Urea Sensitization Caused by Separation of *Helicobacter pylori* RNA Polymerase  $\beta$  and  $\beta'$  Subunits. *Helicobacter*, 12(2), 103-111.
- Damon, I. K. (2013): Poxviruses. In: Knipe, D. M., Howley, P. M., (eds.): *Fields Virology*, 6th edition. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2160-2184.
- Davison, A. J., Moss, B. (1989a): Structure of Vaccinia Virus Early Promoters. *Journal of Molecular Biology*, 210(4), 749-769.
- Davison, A. J., Moss, B. (1989b): Structure of Vaccinia Virus Late Promoters. *Journal of Molecular Biology*, 210(4), 771-784.
- de Hoon, M. J., Makita, Y., Nakai, K., Miyano, S. (2005): Prediction of Transcriptional Terminators in *Bacillus subtilis* and Related Species. *PLoS Computational Biology*, 1(3), e25.
- De la Peña, M., Kyrieleis, O. J., Cusack, S. (2007): Structural insights into the mechanism and evolution of the vaccinia virus mRNA cap N7 methyl-transferase. *The EMBO Journal*, 26(23), 4913-4925.
- de Magistris, L., Stunnenberg, H. G. (1988): Cis-acting sequences affecting the length of the poly(A) head of vaccinia virus late transcripts. *Nucleic Acids Research*, 16(8), 3141-3156.
- De Silva, F. S., Lewis, W., Berglund, P., Koonin, E. V., Moss, B. (2007): Poxvirus DNA primase. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), 18724-18729.
- De Silva, F. S., Moss, B. (2005): Origin-independent plasmid replication occurs in vaccinia virus cytoplasmic factories and requires all five known poxvirus replication factors. *Virology Journal*, 2, 23.

- DeLange, A. M., McFadden, G. (1986): Sequence-nonspecific replication of transfected plasmid DNA in poxvirus-infected cells. *Proceedings of the National Academy of Sciences of the United States of America*, 83(3), 614-618.
- Dellis, S., Strickland, K. C., McCrary, W. J., Patel, A., Stocum, E., Wright, C. F. (2004): Protein interactions among the vaccinia virus late transcription factors. *Virology*, 329(2), 328-336.
- DeMasi, J., Traktman, P. (2000): Clustered Charge-to-Alanine Mutagenesis of the Vaccinia Virus H5 Gene: Isolation of a Dominant, Temperature-Sensitive Mutant with a Profound Defect in Morphogenesis. *Journal of Virology*, 74(5), 2393-2405.
- Deng, L., Beigelman, L., Matulic-Adamic, J., Karpeisky, A., Gershon, P. D. (1997): Specific Recognition of an rU2-N15-rU Motif by VP55, the Vaccinia Virus Poly(A) Polymerase Catalytic Subunit. *Journal of Biological Chemistry*, 272(50), 31542-31552.
- Deng, L., Gershon, P. D. (1997): Interplay of two uridylate-specific RNA binding sites in the translocation of poly(A) polymerase from vaccinia virus. *The EMBO Journal*, 16(5), 1103-1113.
- Deng, L., Hagler, J., Shuman, S. (1996): Factor-dependent Release of Nascent RNA by Ternary Complexes of Vaccinia RNA Polymerase. *Journal of Biological Chemistry*, 271(32), 19556-19562.
- Deng, L., Shuman, S. (1996): An ATPase Component of the Transcription Elongation Complex Is Required for Factor-dependent Transcription Termination by Vaccinia RNA Polymerase. *Journal of Biological Chemistry*, 271(46), 29386-29392.
- Deng, L., Shuman, S. (1997): Elongation Properties of Vaccinia Virus RNA Polymerase: Pausing, Slippage, 3' End Addition, and Termination Site Choice. *Biochemistry*, 36(50), 15892-15899.
- Deng, L., Shuman, S. (1998): Vaccinia NPH-I, a DEXH-box ATPase, is the energy coupling factor for mRNA transcription termination. *Genes & Development*, 12(4), 538-546.
- Dhungel, P., Cao, S., Yang, Z. L. (2017): The 5'-poly(A) leader of poxvirus mRNA confers a translational advantage that can be achieved in cells with impaired cap-dependent translation. *PLoS Pathogens*, 13(8), e1006602.
- Du, S., Traktman, P. (1996): Vaccinia virus DNA replication: Two hundred base pairs of telomeric sequence confer optimal replication efficiency on minichromosome templates. *Proceedings of the National Academy of Sciences of the United States of America*, 93(18), 9693-9698.
- Dürr, H., Körner, C., Müller, M., Hickmann, V., Hopfner, K. P. (2005): X-ray Structures of the Sulfolobus solfataricus SWI2/SNF2 ATPase Core and Its Complex with DNA. *Cell*, 121(3), 363-373.
- Earl, P. L., Hügin, A. W., Moss, B. (1990): Removal of Cryptic Poxvirus Transcription Termination Signals from the Human Immunodeficiency Virus Type 1 Envelope Gene Enhances Expression and Immunogenicity of a Recombinant Vaccinia Virus. *Journal of Virology*, 64(5), 2448-2451.
- Easterbrook, K. B. (1966): Controlled Degradation of Vaccinia Virions in Vitro: an Electron Microscopic Study. *Journal of Ultrastructure Research*, 14(5), 484-496.
- Eun, H.-M. (1996): RNA Polymerases. *Enzymology Primer for Recombinant DNA Technology*. San Diego: Academic Press, 491-565.
- Evans, E., Klemperer, N., Ghosh, R., Traktman, P. (1995): The Vaccinia Virus D5 Protein, Which Is Required for DNA Replication, Is a Nucleic Acid-Independent Nucleoside Triphosphatase. *Journal of Virology*, 69(9), 5353-5361.
- Fabre, E., Jeudy, S., Santini, S., Legendre, M., Trauchessec, M., Couté, Y., Claverie, J. M., Abergel, C. (2017): Noumeavirus replication relies on a transient remote control of the host nucleus. *Nature Communications*, 8, 15087.
- Fritsch, T. E., Siqueira, F. M., Schrank, I. S. (2015): Intrinsic terminators in Mycoplasma hyopneumoniae transcription. *BMC Genomics*, 16, 273.
- Fukuda, K., Maebuchi, M., Takata, H., Gunge, N. (1997): The linear plasmid pDHL1 from Debaryomyces hansenii encodes a protein highly homologous to the pGKL1-plasmid DNA polymerase. *Yeast*, 13(7), 613-620.
- Fukuhara, H. (1995): Linear DNA plasmids of yeasts. *FEMS Microbiology Letters*, 131(1), 1-9.
- Garcés, J., Masternak, K., Kunz, B., Wittek, R. (1993): Reactivation of Transcription from a Vaccinia Virus Early Promoter Late in Infection. *Journal of Virology*, 67(9), 5394-5401.
- Garcia, A. D., Moss, B. (2001): Repression of Vaccinia Virus Holliday Junction Resolvase Inhibits Processing of Viral DNA into Unit-Length Genomes. *Journal of Virology*, 75(14), 6460-6471.
- Garon, C. F., Barbosa, E., Moss, B. (1978): Visualization of an inverted terminal repetition in vaccinia virus DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 75(10), 4863-4867.
- Gaylord, W. H., Melnick, J. L. (1953): Intracellular Forms of Pox Viruses as Shown by the Electron Microscope (Vaccinia, Ectromelia, Molluscum Contagiosum). *Journal of Experimental Medicine*, 98(2), 157-172.
- Gershon, P. D., Ahn, B. Y., Garfield, M., Moss, B. (1991): Poly(A) Polymerase and a Dissociable Polyadenylation Stimulatory Factor Encoded by Vaccinia Virus. *Cell*, 66(6), 1269-1278.
- Gershon, P. D., Moss, B. (1990): Early transcription factor subunits are encoded by vaccinia virus late genes. *Proceedings of the National Academy of Sciences of the United States of America*, 87(11), 4401-4405.



- Gershon, P. D., Moss, B. (1992): Transition from rapid processive to slow nonprocessive polyadenylation by vaccinia virus poly(A) polymerase catalytic subunit is regulated by the net length of the poly(A) tail. *Genes & Development*, 6(8), 1575-1586.
- Gershon, P. D., Moss, B. (1993a): Stimulation of Poly(A) Tail Elongation by the VP39 Subunit of the Vaccinia Virus-encoded Poly(A) Polymerase. *Journal of Biological Chemistry*, 268(3), 2203-2210.
- Gershon, P. D., Moss, B. (1993b): Uridylate-containing RNA sequences determine specificity for binding and polyadenylation by the catalytic subunit of vaccinia virus poly(A) polymerase. *The EMBO Journal*, 12(12), 4705-4714.
- Gershon, P. D., Shi, X. N., Hodel, A. E. (1998): Evidence That the RNA Methylation and Poly(A) Polymerase Stimulatory Activities of Vaccinia Virus Protein VP39 Do Not Impinge upon One Another. *Virology*, 246(2), 253-265.
- Gershowitz, A., Boone, R. F., Moss, B. (1978): Multiple Roles for ATP in the Synthesis and Processing of mRNA by Vaccinia Virus: Specific Inhibitory Effects of Adenosine ( $\beta,\gamma$ -Imido)Triphosphate. *Journal of Virology*, 27(2), 399-408.
- Gershowitz, A., Moss, B. (1979): Abortive Transcription Products of Vaccinia Virus Are Guanylated, Methylated, and Polyadenylated. *Journal of Virology*, 31(3), 849-853.
- Ghosh, A., Lima, C. D. (2010): Enzymology of RNA cap synthesis. *Wiley Interdisciplinary Reviews. RNA*, 1(1), 152-172.
- Gierman, T. M., Frederickson, R. M., Sonenberg, N., Pickup, D. J. (1992): The Eukaryotic Translation initiation Factor 4E Is Not Modified during the Course of Vaccinia Virus Replication. *Virology*, 188(2), 934-937.
- Gilbert, W. V., Zhou, K., Butler, T. K., Doudna, J. A. (2007): Cap-Independent Translation Is Required for Starvation-Induced Differentiation in Yeast. *Science*, 317(5842), 1224-1227.
- Goebel, S. J., Johnson, G. P., Perkus, M. E., Davis, S. W., Winslow, J. P., Paoletti, E. (1990): The Complete DNA Sequence of Vaccinia Virus. *Virology*, 179(1), 247-266.
- Golini, F., Kates, J. R. (1985): A Soluble Transcription System Derived from Purified Vaccinia Virions. *Journal of Virology*, 53(1), 205-213.
- Gong, C., Shuman, S. (2003): Mapping the active site of vaccinia virus RNA triphosphatase. *Virology*, 309(1), 125-134.
- Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., Blinov, V. M. (1989): Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Research*, 17(12), 4713-4730.
- Gordon, B. R. G., Li, Y., Wang, L., Sintsova, A., van Bakel, H., Tian, S., Navarre, W. W., Xia, B., Liu, J. (2010): Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America*, 107(11), 5154-5159.
- Goyer, C., Altmann, M., Lee, H. S., Blanc, A., Deshmukh, M., Woolford, J. L., Jr., Trachsel, H., Sonenberg, N. (1993): TIF4631 and TIF4632: Two Yeast Genes Encoding the High-Molecular-Weight Subunits of the Cap-Binding Protein Complex (Eukaryotic Initiation Factor 4F) Contain an RNA Recognition Motif-Like Sequence and Carry Out an Essential Function. *Molecular and Cellular Biology*, 13(8), 4860-4874.
- Gross, C. H., Shuman, S. (1996a): Vaccinia Virions Lacking the RNA Helicase Nucleoside Triphosphate Phosphohydrolase II Are Defective in Early Transcription. *Journal of Virology*, 70(12), 8549-8557.
- Gross, C. H., Shuman, S. (1996b): Vaccinia Virus RNA Helicase: Nucleic Acid Specificity in Duplex Unwinding. *Journal of Virology*, 70(4), 2615-2619.
- Gross, C. H., Shuman, S. (1998): RNA 5'-Triphosphatase, Nucleoside Triphosphatase, and Guanylyltransferase Activities of Baculovirus LEF-4 Protein. *Journal of Virology*, 72(12), 10020-10028.
- Guarino, L. A., Xu, B., Jin, J., Dong, W. (1998): A Virus-Encoded RNA Polymerase Purified from Baculovirus-Infected Cells. *Journal of Virology*, 72(10), 7985-7991.
- Gudkov, A. T., Ozerova, M. V., Shiryaev, V. M., Spirin, A. S. (2005): 5'-poly(A) Sequence as an Effective Leader for Translation in Eukaryotic Cell-Free Systems. *Biotechnology and Bioengineering*, 91(4), 468-473.
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O. (2010): New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307-321.
- Güldener, U., Heck, S., Fielder, T., Beinhauer, J., Hegemann, J. H. (1996): A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Research*, 24(13), 2519-2524.
- Gunasinghe, S. K., Hubbs, A. E., Wright, C. F. (1998): A Vaccinia Virus Late Transcription Factor with Biochemical and Molecular Identity to a Human Cellular Protein. *Journal of Biological Chemistry*, 273(42), 27524-27530.
- Gunge, N. (1983): Yeast DNA Plasmids. *Annual Review of Microbiology*, 37, 253-276.
- Gunge, N., Murata, K., Sakaguchi, K. (1982): Transformation of Saccharomyces cerevisiae with linear DNA killer plasmids from Kluyveromyces lactis. *Journal of Bacteriology*, 151(1), 462-464.

- Gunge, N., Tamaru, A., Ozawa, F., Sakaguchi, K. (1981): Isolation and characterization of linear deoxyribonucleic acid plasmids from *Kluyveromyces lactis* and the plasmid-associated killer character. *Journal of Bacteriology*, 145(1), 382-390.
- Gunge, N., Yamane, C. (1984): Incompatibility of linear DNA killer plasmids pGKL1 and pGKL2 from *Kluyveromyces lactis* with mitochondrial DNA from *Saccharomyces cerevisiae*. *Journal of Bacteriology*, 159(2), 533-539.
- Gusarov, I., Nudler, E. (1999): The Mechanism of Intrinsic Transcription Termination. *Molecular Cell*, 3(4), 495-504.
- Hagler, J., Luo, Y., Shuman, S. (1994): Factor-dependent Transcription Termination by Vaccinia RNA Polymerase - Kinetic Coupling and Requirement for ATP Hydrolysis. *Journal of Biological Chemistry*, 269(13), 10050-10060.
- Hagler, J., Shuman, S. (1992a): A freeze-frame view of eukaryotic transcription during elongation and capping of nascent mRNA. *Science*, 255(5047), 983-986.
- Hagler, J., Shuman, S. (1992b): Structural analysis of ternary complexes of vaccinia RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 89(21), 10099-10103.
- Hagler, J., Shuman, S. (1992c): Ternary Complex Formation by Vaccinia Virus RNA Polymerase at an Early Viral Promoter: Analysis by Native Gel Electrophoresis. *Journal of Virology*, 66(5), 2982-2989.
- Hagler, J., Shuman, S. (1993): Nascent RNA Cleavage by Purified Ternary Complexes of Vaccinia RNA Polymerase. *Journal of Biological Chemistry*, 268(3), 2166-2173.
- Hall, T. A. (1999): BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98.
- Hänggi, M., Bannwarth, W., Stunnenberg, H. G. (1986): Conserved TAAAT motif in vaccinia virus late promoters: overlapping TATA box and site of transcription initiation. *The EMBO Journal*, 5(5), 1071-1076.
- Harford, C. G., Hamlin, A., Rieders, E. (1966): ELECTRON MICROSCOPIC AUTORADIOGRAPHY OF DNA SYNTHESIS IN CELLS INFECTED WITH VACCINIA VIRUS. *Experimental Cell Research*, 42(1), 50-57.
- Harris, N., Rosales, R., Moss, B. (1993): Transcription initiation factor activity of vaccinia virus capping enzyme is independent of mRNA guanylation. *Proceedings of the National Academy of Sciences of the United States of America*, 90(7), 2860-2864.
- Hassett, D. E., Lewis, J. I., Xing, X., DeLange, L., Condit, R. C. (1997): Analysis of a Temperature-Sensitive Vaccinia Virus Mutant in the Viral mRNA Capping Enzyme Isolated by Clustered Charge-to-Alanine Mutagenesis and Transient Dominant Selection. *Virology*, 238(2), 391-409.
- Hausmann, S., Vivares, C. P., Shuman, S. (2002): Characterization of the mRNA Capping Apparatus of the Microsporidian Parasite *Encephalitozoon cuniculi*. *Journal of Biological Chemistry*, 277(1), 96-103.
- Hausmann, S., Zheng, S., Fabrega, C., Schneller, S. W., Lima, C. D., Shuman, S. (2005): *Encephalitozoon cuniculi* mRNA Cap (Guanine N-7) Methyltransferase: methyl acceptor specificity, inhibition by S-adenosylmethionine analogs, and structure-guided mutational analysis. *Journal of Biological Chemistry*, 280(21), 20404-20412.
- Heilmann, H. J., Magert, H. J., Gassen, H. G. (1988): Identification and isolation of glucose dehydrogenase genes of *Bacillus megaterium* M1286 and their expression in *Escherichia coli*. *European Journal of Biochemistry*, 174(3), 485-490.
- Heuser, J. (2005): Deep-etch EM reveals that the early poxvirus envelope is a single membrane bilayer stabilized by a geodetic "honeycomb" surface coat. *The Journal of Cell Biology*, 169(2), 269-283.
- Higman, M. A., Bourgeois, N., Niles, E. G. (1992): The Vaccinia Virus mRNA (Guanine-N7)-methyltransferase Requires Both Subunits of the mRNA Capping Enzyme for Activity. *Journal of Biological Chemistry*, 267(23), 16430-16437.
- Higman, M. A., Christen, L. A., Niles, E. G. (1994): The mRNA (Guanine-7)-methyltransferase Domain of the Vaccinia Virus mRNA Capping Enzyme: Expression in *Escherichia coli* and structural and kinetic comparison to the intact capping enzyme. *Journal of Biological Chemistry*, 269(21), 14974-14981.
- Hindman, R., Gollnick, P. (2016): Nucleoside Triphosphate Phosphohydrolase I (NPH I) Functions as a 5' to 3' Translocase in Transcription Termination of Vaccinia Early Genes. *The Journal of Biological Chemistry*, 291(28), 14826-14838.
- Hirschmann, P., Vos, J. C., Stunnenberg, H. G. (1990): Mutational Analysis of a Vaccinia Virus Intermediate Promoter In Vivo and In Vitro. *Journal of Virology*, 64(12), 6063-6069.
- Hirtreiter, A., Damsma, G. E., Cheung, A. C. M., Klose, D., Grohmann, D., Vojnic, E., Martin, A. C. R., Cramer, P., Werner, F. (2010): Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. *Nucleic Acids Research*, 38(12), 4040-4051.
- Hishinuma, F., Hirai, K. (1991): Genome organization of the linear plasmid, pSKL, isolated from *Saccharomyces kluyveri*. *Molecular and General Genetics*, 226(1-2), 97-106.
- Ho, C. K., Martins, A., Shuman, S. (2000): A Yeast-Based Genetic System for Functional Analysis of Viral mRNA Capping Enzymes. *Journal of Virology*, 74(12), 5486-5494.
- Hollinshead, M., Rodger, G., Van Eijl, H., Law, M., Hollinshead, R., Vaux, D. J. T., Smith, G. L. (2001): Vaccinia virus utilizes microtubules for movement to the cell surface. *Journal of Cell Biology*, 154(2), 389-402.

- Holm, L., Rosenström, P. (2010): Dali server: conservation mapping in 3D. *Nucleic Acids Research*, 38(Web Server Issue), W545-W549.
- Holowczak, J. A. (1972): Uncoating of Poxviruses I. Detection and Characterization of Subviral Particles in the Uncoating Process. *Virology*, 50(1), 216-232.
- Hooda-Dhingra, U., Thompson, C. L., Condit, R. C. (1989): Detailed Phenotypic Characterization of Five Temperature-Sensitive Mutants in the 22- and 147-Kilodalton Subunits of Vaccinia Virus DNA-Dependent RNA Polymerase. *Journal of Virology*, 63(2), 714-729.
- Howard, S. T., Ray, C. A., Patel, D. D., Antczak, J. B., Pickup, D. J. (1999): A 43-Nucleotide RNA Cis-Acting Element Governs the Site-Specific Formation of the 3' End of a Poxvirus Late mRNA. *Virology*, 255(1), 190-204.
- Hubbs, A. E., Wright, C. F. (1996): The A2L Intermediate Gene Product Is Required for In Vitro Transcription from a Vaccinia Virus Late Promoter. *Journal of Virology*, 70(1), 327-331.
- Hyun, J. K., Accurso, C., Hijnen, M., Schult, P., Pettikiriachchi, A., Mitra, A. K., Coulibaly, F. (2011): Membrane Remodeling by the Double-Barrel Scaffolding Protein of Poxvirus. *PLoS Pathogens*, 7(9), e1002239.
- Hyun, J. K., Coulibaly, F., Turner, A. P., Baker, E. N., Mercer, A. A., Mitra, A. K. (2007): The Structure of a Putative Scaffolding Protein of Immature Poxvirus Particles as Determined by Electron Microscopy Suggests Similarity with Capsid Proteins of Large Icosahedral DNA Viruses. *Journal of Virology*, 81(20), 11075-11083.
- Chiang, S. Y., Welch, J., Rauscher, F. J., III, Beerman, T. A. (1994): Effects of Minor Groove Binding Drugs on the Interaction of TATA Box Binding Protein and TFIIA with DNA. *Biochemistry*, 33(23), 7033-7040.
- Chlanda, P., Carbajal, M. A., Cyrklaff, M., Griffiths, G., Krijnse-Locker, J. (2009): Membrane Rupture Generates Single Open Membrane Sheets during Vaccinia Virus Assembly. *Cell Host & Microbe*, 6(1), 81-90.
- Cho, E.-J., Rodriguez, C. R., Takagi, T., Buratowski, S. (1998): Allosteric interactions between capping enzyme subunits and the RNA polymerase II carboxy-terminal domain. *Genes & Development*, 12(22), 3482-3487.
- Christen, L. A., Piacente, S., Mohamed, M. R., Niles, E. G. (2008): Vaccinia virus early gene transcription termination factors VTF and Rap94 interact with the U9 termination motif in the nascent RNA in a transcription ternary complex. *Virology*, 376(1), 225-235.
- Christen, L. M., Sanders, M., Wiler, C., Niles, E. G. (1998): Vaccinia virus nucleoside triphosphate phosphohydrolase I is an essential viral early gene transcription termination factor. *Virology*, 245(2), 360-371.
- Chung, C. S., Chen, C. H., Ho, M. Y., Huang, C. Y., Liao, C. L., Chang, W. (2006): Vaccinia Virus Proteome: Identification of Proteins in Vaccinia Virus Intracellular Mature Virion Particles. *Journal of Virology*, 80(5), 2127-2140.
- İnce, İ. A., Pijlman, G. P., Vlak, J. M., van Oers, M. M. (2017): Hairpin structures with conserved sequence motifs determine the 3' ends of non-polyadenylated invertebrate iridovirus transcripts. *Virology*, 511, 344-353.
- Ingham, C. J., Hunter, I. S., Smith, M. C. M. (1995): Rho-independent terminators without 3' poly-U tails from the early region of actinophage  $\phi$ C31. *Nucleic Acids Research*, 23(3), 370-376.
- Ink, B. S., Pickup, D. J. (1990): Vaccinia virus directs the synthesis of early mRNAs containing 5' poly(A) sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 87(4), 1536-1540.
- Ishii, K., Moss, B. (2002): Mapping Interaction Sites of the A20R Protein Component of the Vaccinia Virus DNA Replication Complex. *Virology*, 303(2), 232-239.
- Iyer, L. M., Aravind, L. (2012): Insights from the architecture of the bacterial transcription apparatus. *Journal of Structural Biology*, 179(3), 299-319.
- Iyer, L. M., Aravind, L., Koonin, E. V. (2001): Common origin of four diverse families of large eukaryotic DNA viruses. *Journal of Virology*, 75(23), 11720-11734.
- Iyer, L. M., Balaji, S., Koonin, E. V., Aravind, L. (2006): Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Research*, 117(1), 156-184.
- Iyer, L. M., Koonin, E. V., Aravind, L. (2003): Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. *BMC Structural Biology*, 3, 1.
- Jeske, S., Meinhardt, F. (2006): Autonomous cytoplasmic linear plasmid pPac1-1 of *Pichia acaciae*: molecular structure and expression studies. *Yeast*, 23(6), 479-486.
- Jeske, S., Meinhardt, F., Klassen, R. (2007): Extranuclear Inheritance: Virus-Like DNA-Elements in Yeast. In: Esser, K., Löttge, U., Beyschlag, W., Murata, J., (eds.): *Progress in Botany*, 68. Springer-Verlag, Berlin, Heidelberg, 98-129.
- Jeske, S., Tiggemann, M., Meinhardt, F. (2006): Yeast autonomous linear plasmid pGKL2: ORF9 is an actively transcribed essential gene with multiple transcription start points. *FEMS Microbiology Letters*, 255(2), 321-327.
- Jha, S., Rollins, M. G., Fuchs, G., Procter, D. J., Hall, E. A., Cozzolino, K., Sarnow, P., Savas, J. N., Walsh, D. (2017): Trans-kingdom mimicry underlies ribosome customization by a poxvirus kinase. *Nature*, 546(7660), 651-655.

- Jin, J., Dong, W., Guarino, L. A. (1998): The LEF-4 Subunit of Baculovirus RNA Polymerase Has RNA 5'-Triphosphatase and ATPase Activities. *Journal of Virology*, 72(12), 10011-10019.
- Johnson, L., Gupta, A. T. K., Ghafoor, A., Akin, D., Bashir, R. (2006): Characterization of vaccinia virus particles using microscale silicon cantilever resonators and atomic force microscopy. *Sensors and Actuators B: Chemical*, 115(1), 189-197.
- Johnson, L., Liu, S., Gershon, P. D. (2004): Molecular Flexibility and Discontinuous Translocation of a Non-templated Polymerase. *Journal of Molecular Biology*, 337(4), 843-856.
- Jones, E. V., Moss, B. (1984): Mapping of the Vaccinia Virus DNA Polymerase Gene by Marker Rescue and Cell-Free Translation of Selected RNA. *Journal of Virology*, 49(1), 72-77.
- Jong, A. Y., Campbell, J. L. (1986): Isolation of the gene encoding yeast single-stranded nucleic acid binding protein 1. *Proceedings of the National Academy of Sciences of the United States of America*, 83(4), 877-881.
- Jungwirth, C., Launer, J. (1968): Effect of poxvirus infection on host cell deoxyribonucleic acid synthesis. *Journal of Virology*, 2(5), 401-408.
- Kadaré, G., Haenni, A. L. (1997): Virus-Encoded RNA Helicases. *Journal of Virology*, 71(4), 2583-2590.
- Kämper, J., Esser, K., Gunge, N., Meinhardt, F. (1991): Heterologous gene expression on the linear DNA killer plasmid from *Kluyveromyces lactis*. *Current Genetics*, 19(2), 109-118.
- Kämper, J., Meinhardt, F., Gunge, N., Esser, K. (1989a): In vivo construction of linear vectors based on killer plasmids from *Kluyveromyces lactis*: selection of a nuclear gene results in attachment of telomeres. *Molecular and Cellular Biology*, 9(9), 3931-3937.
- Kämper, J., Meinhardt, F., Gunge, N., Esser, K. (1989b): New recombinant linear DNA-elements derived from *Kluyveromyces lactis* killer plasmids. *Nucleic Acids Research*, 17(4), 1781.
- Kapitonov, V. V., Jurka, J. (2006): Self-synthesizing DNA transposons in eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12), 4540-4545.
- Kast, A., Voges, R., Schroth, M., Schaffrath, R., Klassen, R., Meinhardt, F. (2015): Autoselection of Cytoplasmic Yeast Virus Like Elements Encoding Toxin/Antitoxin Systems Involves a Nuclear Barrier for Immunity Gene Expression. *PLoS Genetics*, 11(5), e1005005.
- Kates, J., Beeson, J. (1970a): Ribonucleic Acid Synthesis in Vaccinia Virus I. The Mechanism of Synthesis and Release of RNA in Vaccinia Cores. *Journal of Molecular Biology*, 50(1), 1-18.
- Kates, J., Beeson, J. (1970b): Ribonucleic Acid Synthesis in Vaccinia Virus II. Synthesis of Polyriboadenylic Acid. *Journal of Molecular Biology*, 50(1), 19-33.
- Kates, J. R., McAuslan, B. R. (1967): MESSENGER RNA SYNTHESIS BY A "COATED" VIRAL GENOME. *Proceedings of the National Academy of Sciences of the United States of America*, 57(2), 314-320.
- Katsafanas, G. C., Moss, B. (2004): Vaccinia Virus Intermediate Stage Transcription Is Complemented by Ras-GTPase-activating Protein SH3 Domain-binding Protein (G3BP) and Cytoplasmic Activation/Proliferation-associated Protein (p137) Individually or as a Heterodimer. *Journal of Biological Chemistry*, 279(50), 52210-52217.
- Katsafanas, G. C., Moss, B. (2007): Colocalization of Transcription and Translation within Cytoplasmic Poxvirus Factories Coordinates Viral Expression and Subjugates Host Functions. *Cell Host & Microbe*, 2(4), 221-228.
- Kay, N. E., Bainbridge, T. W., Condit, R. C., Bubb, M. R., Judd, R. E., Venkatakrishnan, B., McKenna, R., D'Costa, S. M. (2013): Biochemical and Biophysical Properties of a Putative Hub Protein Expressed by Vaccinia Virus. *Journal of Biological Chemistry*, 288(16), 11470-11481.
- Keck, J. G., Baldick, C. J., Jr., Moss, B. (1990): Role of DNA Replication in Vaccinia Virus Gene Expression: A Naked Template Is Required for Transcription of Three Late Trans-Activator Genes. *Cell*, 61(5), 801-809.
- Keck, J. G., Feigenbaum, F., Moss, B. (1993a): Mutational Analysis of a Predicted Zinc-Binding Motif in the 26-Kilodalton Protein Encoded by the Vaccinia Virus A2L Gene: Correlation of Zinc Binding with Late Transcriptional Transactivation Activity. *Journal of Virology*, 67(10), 5749-5753.
- Keck, J. G., Kovacs, G. R., Moss, B. (1993b): Overexpression, Purification, and Late Transcription Factor Activity of the 17-Kilodalton Protein Encoded by the Vaccinia Virus ALL Gene. *Journal of Virology*, 67(10), 5740-5748.
- Keith, J. M., Gershowitz, A., Moss, B. (1980): Dinucleotide Sequences at the 5' Ends of Vaccinia Virus mRNA's Synthesized In Vitro. *Journal of Virology*, 36(2), 601-605.
- Kempken, F., Hermanns, J., Osiewacz, H. D. (1992): Evolution of linear plasmids. *Journal of Molecular Evolution*, 35(6), 502-513.
- Kettenberger, H., Armache, K. J., Cramer, P. (2004): Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Molecular Cell*, 16(6), 955-965.
- Kikuchi, Y., Hirai, K., Gunge, N., Hishinuma, F. (1985): Hairpin plasmid — a novel linear DNA of perfect hairpin structure. *The EMBO Journal*, 4(7), 1881-1886.
- Kikuchi, Y., Hirai, K., Hishinuma, F. (1984): The yeast linear DNA killer plasmids, pGKL1 and pGKL2, possess terminally attached proteins. *Nucleic Acids Research*, 12(14), 5685-5692.
- Kim, D. E., Chivian, D., Baker, D. (2004): Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research*, 32(Web Server Issue), W526-W531.
- Kim, J. H. (2004): Immobilized DNA-binding assay, an approach for in vitro DNA-binding assay. *Analytical Biochemistry*, 334(2), 401-402.

- Kim, J. L., Nikolov, D. B., Burley, S. K. (1993a): Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*, 365(6446), 520-527.
- Kim, Y., Geiger, J. H., Hahn, S., Sigler, P. B. (1993b): Crystal structure of a yeast TBP/TATA-box complex. *Nature*, 365(6446), 512-520.
- Kitada, K., Gunge, N. (1988): Palindrome-hairpin linear plasmids possessing only a part of the ORF1 gene of the yeast killer plasmid pGKL1. *Molecular & General Genetics*, 215(1), 46-52.
- Klassen, R., Jablonowski, D., Schaffrath, R., Meinhardt, F. (2002): Genome organization of the linear *Pichia etchellsii* plasmid pPE1A: evidence for expression of an extracellular chitin-binding protein homologous to the alpha-subunit of the *Kluyveromyces lactis* killer toxin. *Plasmid*, 47(3), 224-233.
- Klassen, R., Meinhardt, F. (2003): Structural and functional analysis of the killer element pPin1-3 from *Pichia inositovora*. *Molecular Genetics and Genomics*, 270(2), 190-199.
- Klassen, R., Meinhardt, F. (2007): Linear Protein-Primed Replicating Plasmids in Eukaryotic Microbes. In: Meinhardt, F., Klassen, R., (eds.): *Microbial Linear Plasmids*, 7. Springer-Verlag, Berlin, Heidelberg, 187-226.
- Klassen, R., Tontsidou, L., Larsen, M., Meinhardt, F. (2001): Genome organization of the linear cytoplasmic element pPE1B from *Pichia etchellsii*. *Yeast*, 18(10), 953-961.
- Knutson, B. A., Broyles, S. S. (2008): Expansion of poxvirus RNA polymerase subunits sharing homology with corresponding subunits of RNA polymerase II. *Virus Genes*, 36(2), 307-311.
- Knutson, B. A., Drennan, M., Liu, X., Broyles, S. S. (2009a): Bidirectional transcriptional promoters in the vaccinia virus genome. *Virology*, 385(1), 198-203.
- Knutson, B. A., Liu, X., Oh, J., Broyles, S. S. (2006): Vaccinia Virus Intermediate and Late Promoter Elements Are Targeted by the TATA-Binding Protein. *Journal of Virology*, 80(14), 6784-6793.
- Knutson, B. A., Oh, J., Broyles, S. S. (2009b): Downregulation of vaccinia virus intermediate and late promoters by host transcription factor YY1. *Journal of General Virology*, 90(Pt 7), 1592-1599.
- Koonin, E. V., Dolja, V. V., Krupović, M. (2015a): Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology*, 479-480, 2-25.
- Koonin, E. V., Krupović, M., Yutin, N. (2015b): Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. *Annals of the New York Academy of Sciences*, 1341, 10-24.
- Koonin, E. V., Senkevich, T. G. (1992): Vaccinia virus encodes four putative DNA and/or RNA helicases distantly related to each other. *Journal of General Virology*, 73(Pt 4), 989-993.
- Koonin, E. V., Yutin, N. (2010): Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. *Intervirology*, 53(5), 284-292.
- Koutmou, K. S., Schuller, A. P., Brunelle, J. L., Radhakrishnan, A., Djuranovic, S., Green, R. (2015): Ribosomes slide on lysine-encoding homopolymeric A stretches. *eLife*, 4, e05534.
- Kovacs, G. R., Moss, B. (1996): The Vaccinia Virus H5R Gene Encodes Late Gene Transcription Factor 4: Purification, Cloning, and Overexpression. *Journal of Virology*, 70(10), 6796-6802.
- Kovacs, G. R., Rosales, R., Keck, J. G., Moss, B. (1994): Modification of the Cascade Model for Regulation of Vaccinia Virus Gene Expression: Purification of a Prereplicative, Late-Stage-Specific Transcription Factor. *Journal of Virology*, 68(5), 3443-3447.
- Krupović, M., Bamford, D. H., Koonin, E. V. (2014): Conservation of major and minor jelly-roll capsid proteins in Polinton (Maverick) transposons suggests that they are bona fide viruses. *Biology Direct*, 9, 6.
- Krupović, M., Koonin, E. V. (2015): Polintons: a hotbed of eukaryotic virus, transposon and plasmid evolution. *Nature Reviews Microbiology*, 13(2), 105-115.
- Kumar, S., Boyle, D. B. (1990): Activity of a fowlpox virus late gene promoter in vaccinia and fowlpox virus recombinants. *Archives of Virology*, 112(3-4), 139-148.
- Künzi, M. S., Traktman, P. (1989): Genetic Evidence for Involvement of Vaccinia Virus DNA-Dependent ATPase I in Intermediate and Late Gene Expression. *Journal of Virology*, 63(9), 3999-4010.
- Kuznedelov, K., Korzheva, N., Mustaev, A., Severinov, K. (2002): Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. *The EMBO Journal*, 21(6), 1369-1378.
- Kwapisz, M., Beckouët, F., Thuriaux, P. (2008): Early evolution of eukaryotic DNA-dependent RNA polymerases. *Trends in Genetics*, 24(5), 211-215.
- Kyrieleis, O. J., Chang, J., De la Peña, M., Shuman, S., Cusack, S. (2014): Crystal Structure of Vaccinia Virus mRNA Capping Enzyme Provides Insights into the Mechanism and Evolution of the Capping Apparatus. *Structure*, 22(3), 452-465.
- Lackner, C. A., Condit, R. C. (2000): Vaccinia Virus Gene A18R DNA Helicase Is a Transcript Release Factor. *Journal of Biological Chemistry*, 275(2), 1485-1494.
- Laliberte, J. P., Moss, B. (2010): Lipid Membranes in Poxvirus Replication. *Viruses*, 2(4), 972-986.
- Lamberte, L. E., Baniulyte, G., Singh, S. S., Stringer, A. M., Bonocora, R. P., Stracy, M., Kapanidis, A. N., Wade, J. T., Grainger, D. C. (2017): Horizontally acquired AT-rich genes in *Escherichia coli* cause toxicity by sequestering RNA polymerase. *Nature Microbiology*, 2, 16249.
- Lane, W. J., Darst, S. A. (2010a): Molecular evolution of multisubunit RNA polymerases: sequence analysis. *Journal of Molecular Biology*, 395(4), 671-685.

- Lane, W. J., Darst, S. A. (2010b): Molecular evolution of multisubunit RNA polymerases: structural analysis. *Journal of Molecular Biology*, 395(4), 686-704.
- Langberg, S. R., Moss, B. (1981): Post-transcriptional modifications of mRNA. Purification and characterization of cap I and cap II RNA (nucleoside-2')-methyltransferases from HeLa cells. *Journal of Biological Chemistry*, 256(19), 10054-10060.
- Larsen, M., Gunge, N., Meinhardt, F. (1998): Kluveromyces lactis killer plasmid pGKL2: evidence for a viral-like capping enzyme encoded by ORF3. *Plasmid*, 40(3), 243-246.
- Larsen, M., Meinhardt, F. (2000): Kluveromyces lactis killer system: identification of a new gene encoded by pGKL2. *Current Genetics*, 38(5), 271-275.
- Latner, D. R., Thompson, J. M., Gershon, P. D., Storrs, C., Condit, R. C. (2002): The Positive Transcription Elongation Factor Activity of the Vaccinia Virus J3 Protein Is Independent from Its (Nucleoside-2'-O-) Methyltransferase and Poly(A) Polymerase Stimulatory Functions. *Virology*, 301(1), 64-80.
- Latner, D. R., Xiang, Y., Lewis, J. I., Condit, J., Condit, R. C. (2000): The Vaccinia Virus Bifunctional Gene J3 (Nucleoside-2'-O-)methyltransferase and Poly(A) Polymerase Stimulatory Factor Is Implicated as a Positive Transcription Elongation Factor by Two Genetic Approaches. *Virology*, 269(2), 345-355.
- Lee-Chen, G. J., Bourgeois, N., Davidson, K., Condit, R. C., Niles, E. G. (1988): Structure of the Transcription Initiation and Termination Sequences of Seven Early Genes in the Vaccinia Virus HindIII D Fragment. *Virology*, 163(1), 64-79.
- Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., Lescot, M., Poirot, O., Bertaux, L., Bruley, C., Couté, Y., Rivkina, E., Abergel, C., Claverie, J. M. (2014): Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proceedings of the National Academy of Sciences of the United States of America* 111(11), 4274-4279.
- Leitão, A. L., Costa, M. C., Enguita, F. J. (2015): Unzippers, Resolvers and Sensors: A Structural and Functional Biochemistry Tale of RNA Helicases. *International Journal of Molecular Sciences*, 16(2), 2269-2293.
- Li, J., Broyles, S. S. (1993a): The DNA-dependent ATPase activity of vaccinia virus early gene transcription factor is essential for its transcription activation function. *Journal of Biological Chemistry*, 268(27), 20016-20021.
- Li, J., Broyles, S. S. (1993b): Recruitment of vaccinia virus RNA polymerase to an early gene promoter by the viral early transcription factor. *Journal of Biological Chemistry*, 268(4), 2773-2780.
- Li, J., Broyles, S. S. (1995): The DNA binding domain of the vaccinia virus early transcription factor small subunit is an extended helicase-like motif. *Nucleic Acids Research*, 23(9), 1590-1596.
- Li, J., Pennington, M. J., Broyles, S. S. (1994): Temperature-Sensitive Mutations in the Gene Encoding the Small Subunit of the Vaccinia Virus Early Transcription Factor Impair Promoter Binding, Transcription Activation, and Packaging of Multiple Virion Components. *Journal of Virology*, 68(4), 2605-2614.
- Liem, J., Liu, J. (2016): Stress Beyond Translation: Poxviruses and More. *Viruses*, 8(6), 169.
- Lima, C. D., Wang, L. K., Shuman, S. (1999): Structure and Mechanism of Yeast RNA Triphosphatase: An Essential Component of the mRNA Capping Apparatus. *Cell*, 99(5), 533-543.
- Lin, Y. C. J., Evans, D. H. (2010): Vaccinia Virus Particles Mix Inefficiently, and in a Way That Would Restrict Viral Recombination, in Coinfected Cells. *Journal of Virology*, 84(5), 2432-2443.
- Love, K. R., Shah, K. A., Whittaker, C. A., Wu, J., Bartlett, M. C., Ma, D. D., Leeson, R. L., Priest, M., Borowsky, J., Young, S. K., Love, J. C. (2016): Comparative genomics and transcriptomics of Pichia pastoris. *BMC Genomics*, 17, 550.
- Lu, C., Bablanian, R. (1996): Characterization of small nontranslated polyadenylated RNAs in vaccinia virus-infected cells. *Proceedings of the National Academy of Sciences of the United States of America*, 93(5), 2037-2042.
- Lu, Z. Q., Li, Y., Que, Q. D., Kutish, G. F., Rock, D. L., VanEtten, J. L. (1996): Analysis of 94 kb of the Chlorella Virus PBCV-1 330-kb Genome: Map Positions 88 to 182. *Virology*, 216(1), 102-123.
- Lucchini, S., Rowley, G., Goldberg, M. D., Hurd, D., Harrison, M., Hinton, J. C. (2006): H-NS Mediates the Silencing of Laterally Acquired Genes in Bacteria. *PLoS Pathogens*, 2(8), e81.
- Luo, Y., Hagler, J., Shuman, S. (1991): Discrete Functional Stages of Vaccinia Virus Early Transcription during a Single Round of RNA Synthesis in Vitro. *Journal of Biological Chemistry*, 266(20), 13303-13310.
- Luo, Y., Mao, X., Deng, L., Cong, P., Shuman, S. (1995): The D1 and D12 subunits are both essential for the transcription termination factor activity of vaccinia virus capping enzyme. *Journal of Virology*, 69(6), 3852-3856.
- Luo, Y., Shuman, S. (1991): Antitermination of Vaccinia Virus Early Transcription: Possible Role of RNA Secondary Structure. *Virology*, 185(1), 432-436.
- Luo, Y., Shuman, S. (1993): RNA Binding Properties of Vaccinia Virus Capping Enzyme. *Journal of Biological Chemistry*, 268(28), 21253-21262.
- Mackett, M., Archard, L. C. (1979): Conservation and Variation in Orthopoxvirus Genome Structure. *Journal of General Virology*, 45(3), 683-701.
- Mahr, A., Roberts, B. E. (1984a): Arrangement of Late RNAs Transcribed from a 7.1-Kilobase EcoRI Vaccinia Virus DNA Fragment. *Journal of Virology*, 49(2), 510-520.
- Mahr, A., Roberts, B. E. (1984b): Organization of Six Early Transcripts Synthesized from a Vaccinia Virus EcoRI DNA Fragment. *Journal of Virology*, 49(2), 497-509.



- Mallardo, M., Schleich, S., Krijnse Locker, J.** (2001): Microtubule-dependent Organization of Vaccinia Virus Core-derived Early mRNAs into Distinct Cytoplasmic Structures. *Molecular Biology of the Cell*, 12(12), 3875-3891.
- Mao, X., Shuman, S.** (1994): Intrinsic RNA (guanine-7) methyltransferase activity of the vaccinia virus capping enzyme D1 subunit is stimulated by the D12 subunit. Identification of amino acid residues in the D1 protein required for subunit association and methyl group transfer. *Journal of Biological Chemistry*, 269(39), 24472-24479.
- Mao, X., Shuman, S.** (1996): Vaccinia Virus mRNA (Guanine-7-)Methyltransferase: Mutational Effects on Cap Methylation and AdoHcy-Dependent Photo-Cross-Linking of the Cap to the Methyl Acceptor Site. *Biochemistry*, 35(21), 6900-6910.
- Mars, M., Beaud, G.** (1987): Characterization of Vaccinia Virus Early Promoters and Evaluation of Their Informational Content. *Journal of Molecular Biology*, 198(4), 619-631.
- Martin, F. H., Tinoco, I., Jr.** (1980): DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Research*, 8(10), 2295-2299.
- Martin, S. A., Moss, B.** (1975): Modification of RNA by mRNA Guanylyltransferase and mRNA(guanine-7)methyltransferase from Vaccinia Virions. *Journal of Biological Chemistry*, 250(24), 9330-9335.
- Martin, S. A., Paoletti, E., Moss, B.** (1975): Purification of mRNA Guanylyltransferase and mRNA(guanine-7-)methyltransferase from Vaccinia Virions. *Journal of Biological Chemistry*, 250(24), 9322-9329.
- Martins, A., Gross, C. H., Shuman, S.** (1999): Mutational Analysis of Vaccinia Virus Nucleoside Triphosphate Phosphohydrolase I, a DNA-Dependent ATPase of the DExH Box Family. *Journal of Virology*, 73(2), 1302-1308.
- Maruyama, K., Sugano, S.** (1994): Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene*, 138(1-2), 171-174.
- Mašek, T., Vopálenský, V., Suchomelová, P., Pospíšek, M.** (2005): Denaturing RNA electrophoresis in TAE agarose gels. *Analytical Biochemistry*, 336(1), 46-50.
- McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Poster, S., Shuman, S., Bentley, D. L., Program, A. E.** (1997): 5'-capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes & Development*, 11(24), 3306-3318.
- McCraith, S., Holtzman, T., Moss, B., Fields, S.** (2000): Genome-wide analysis of vaccinia virus protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 97(9), 4879-4884.
- McMahon, R., Zaborowska, I., Walsh, D.** (2011): Noncytotoxic Inhibition of Viral Infection through eIF4F-Independent Suppression of Translation by 4EGI-1. *Journal of Virology*, 85(2), 853-864.
- McNeel, D. G., Tamanoi, F.** (1991): Terminal region recognition factor 1, a DNA-binding protein recognizing the inverted terminal repeats of the pGK1 linear DNA plasmids. *Proceedings of the National Academy of Sciences of the United States of America*, 88(24), 11398-11402.
- Meis, R. J., Condit, R. C.** (1991): Genetic and Molecular Biological Characterization of a Vaccinia Virus Gene which Renders the Virus Dependent on Isatin- $\beta$ -thiosemicarbazone (IBT). *Virology*, 182(2), 442-454.
- Merchinsky, M.** (1990): Mutational Analysis of the Resolution Sequence of Vaccinia Virus DNA: Essential Sequence Consists of Two Separate AT-Rich Regions Highly Conserved among Poxviruses. *Journal of Virology*, 64(10), 5029-5035.
- Metz, D. H., Esteban, M., Danieleescu, G.** (1975): The Formation of Virus Polyribosomes in L Cells Infected with Vaccinia Virus. *Journal of General Virology*, 27(2), 181-195.
- Milligan, J. F., Uhlenbeck, O. C.** (1989): Synthesis of Small RNAs Using T7 RNA Polymerase. *Methods in Enzymology*, 180, 51-62.
- Millns, A. K., Carpenter, M. S., Delange, A. M.** (1994): The Vaccinia Virus-Encoded Uracil DNA Glycosylase Has an Essential Role in Viral DNA Replication. *Virology*, 198(2), 504-513.
- Miner, J. N., Weinrich, S. L., Hruby, D. E.** (1988): Molecular Dissection of cis-Acting Regulatory Elements from 5'-Proximal Regions of a Vaccinia Virus Late Gene Cluster. *Journal of Virology*, 62(1), 297-304.
- Minnigan, H., Moyer, R. W.** (1985): Intracellular Location of Rabbit Poxvirus Nucleic Acid Within Infected Cells as Determined by In Situ Hybridization. *Journal of Virology*, 55(3), 634-643.
- Mirzakhanyan, Y., Gershon, P. D.** (2017): Multisubunit DNA-Dependent RNA Polymerases from Vaccinia Virus and Other Nucleocytoplasmic Large-DNA Viruses: Impressions from the Age of Structure. *Microbiology and Molecular Biology Reviews*, 81(3), e00010-e00017.
- Mitra, A., Angamuthu, K., Jayashree, H. V., Nagaraja, V.** (2009): Occurrence, divergence and evolution of intrinsic terminators across Eubacteria. *Genomics*, 94(2), 110-116.
- Mohamed, M. R., Niles, E. G.** (2000): Interaction between Nucleoside Triphosphate Phosphohydrolase I and the H4L Subunit of the Viral RNA Polymerase Is Required for Vaccinia Virus Early Gene Transcript Release. *Journal of Biological Chemistry*, 275(33), 25798-25804.
- Mohamed, M. R., Niles, E. G.** (2001): The Viral RNA Polymerase H4L Subunit Is Required for Vaccinia Virus Early Gene Transcription Termination. *Journal of Biological Chemistry*, 276(23), 20758-20765.

- Mohamed, M. R., Niles, E. G.** (2003a): UUUUUNU Oligonucleotide Stimulation of Vaccinia Virus Early Gene Transcription Termination, in Trans. *Journal of Biological Chemistry*, 278(14), 11794-11801.
- Mohamed, M. R., Niles, E. G.** (2003b): UUUUUNU Stimulation of Vaccinia Virus Early Gene Transcription Termination. Oligonucleotide sequence and structural requirements for stimulation of premature termination in vitro. *Journal of Biological Chemistry*, 278(41), 39534-39541.
- Mohamed, M. R., Niles, E. G.** (2004): UUUUUNU oligonucleotide inhibition of RNA synthesis in vaccinia virus cores. *Virology*, 324(2), 493-500.
- Moore, J. B., Smith, G. L.** (1992): Steroid hormone synthesis by a vaccinia enzyme: a new type of virus virulence factor. *The EMBO Journal*, 11(5), 1973-1980.
- Morgan, C., Ellison, S. A., Rose, H. M., Moore, D. H.** (1954): Structure and Development of Viruses Observed in the Electron Microscope. II. Vaccinia and Fowl Pox Viruses. *Journal of Experimental Medicine*, 100(3), 301-310.
- Morgan, J. R., Cohen, L. K., Roberts, B. E.** (1984): Identification of the DNA Sequences Encoding the Large Subunit of the mRNA-Capping Enzyme of Vaccinia Virus. *Journal of Virology*, 52(1), 206-14.
- Moss, B.** (2013a): Poxvirus DNA replication. *Cold Spring Harbor Perspectives in Biology*, 5(9), a010199.
- Moss, B.** (2013b): Poxviruses. In: Knipe, D. M., Howley, P. M., (eds.): *Fields Virology*, 6th edition. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2129-2159.
- Moss, B.** (2015): Poxvirus membrane biogenesis. *Virology*, 479-480, 619-626.
- Moss, B.** (2016): Membrane fusion during poxvirus entry. *Seminars in Cell & Developmental Biology*, 60, 89-96.
- Moss, B., Rosenblum, E. N., Gershowitz, A.** (1975): Characterization of a Polyriboadenylate Polymerase from Vaccinia Virions. *Journal of Biological Chemistry*, 250(12), 4722-4729.
- Moss, B., Winters, E., Jones, E.** (1983): Replication of vaccinia virus. In: Cozzarelli, N., (ed.): *Mechanics of DNA replication and recombination*. New York, 449-461.
- Moure, C. M., Bowman, B. R., Gershon, P. D., Quijcho, F. A.** (2006): Crystal Structures of the Vaccinia Virus Polyadenylate Polymerase Heterodimer: Insights into ATP Selectivity and Processivity. *Molecular Cell*, 22(3), 339-349.
- Moyer, R. W., Graves, R. L.** (1981): The Mechanism of Cytoplasmic Orthopoxvirus DNA Replication. *Cell*, 27(2 Pt 1), 391-401.
- Mulder, J., Robertson, M. E. M., Seamons, R. A., Belsham, G. J.** (1998): Vaccinia Virus Protein Synthesis Has a Low Requirement for the Intact Translation Initiation Factor eIF4F, the Cap-Binding Complex, within Infected Cells. *Journal of Virology*, 72(11), 8813-8819.
- Myette, J. R., Niles, E. G.** (1996a): Domain Structure of the Vaccinia Virus mRNA Capping Enzyme. Expression in Escherichia coli of a subdomain possessing the RNA 5'-triphosphatase and guanylyltransferase activities and a kinetic comparison to the full-size enzyme. *Journal of Biological Chemistry*, 271(20), 11936-11944.
- Myette, J. R., Niles, E. G.** (1996b): Characterization of the Vaccinia Virus RNA 5'-Triphosphatase and Nucleoside Triphosphate Phosphohydrolase Activities. Demonstration that both activities are carried out at the same active site. *Journal of Biological Chemistry*, 271(20), 11945-11952.
- Naji, S., Bertero, M. G., Spitalny, P., Cramer, P., Thomm, M.** (2008): Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. *Nucleic Acids Research*, 36(2), 676-687.
- Nandhagopal, N., Simpson, A. A., Gurnon, J. R., Yan, X., Baker, T. S., Graves, M. V., Van Etten, J. L., Rossmann, M. G.** (2002): The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. *Proceedings of the National Academy of Sciences of the United States of America*, 99(23), 14758-14763.
- Navarre, W. W., Porwollik, S., Wang, Y., McClelland, M., Rosen, H., Libby, S. J., Fang, F. C.** (2006): Selective Silencing of Foreign DNA with Low GC Content by the H-NS Protein in Salmonella. *Science*, 313(5784), 236-238.
- Nevins, J. R., Joklik, W. K.** (1975): Poly(A) Sequences of Vaccinia Virus Messenger RNA: Nature, Mode of Addition and Function during Translation in Vitro and in Vivo. *Virology*, 63(1), 1-14.
- Niles, E. G., Lee-Chen, G. J., Shuman, S., Moss, B., Broyles, S. S.** (1989): Vaccinia Virus Gene D12L Encodes the Small Subunit of the Viral mRNA Capping Enzyme. *Virology*, 172(2), 513-522.
- Novy, K., Kilcher, S., Omasits, U., Bleck, C. K. E., Beerli, C., Vowinckel, J., Martin, C. K., Syedbasha, M., Maiolica, A., White, I., Mercer, J., Wollscheid, B.** (2018): Proteotype profiling unmasks a viral signalling network essential for poxvirus assembly and transcriptional competence. *Nature Microbiology*, 3(5), 588-599.
- Oda, K. I., Joklik, W. K.** (1967): Hybridization and Sedimentation Studies on "Early" and "Late" Vaccinia Messenger RNA. *Journal of Molecular Biology*, 27(3), 395-419.
- Oh, J., Broyles, S. S.** (2005): Host Cell Nuclear Proteins Are Recruited to Cytoplasmic Vaccinia Virus Replication Complexes. *Journal of Virology*, 79(20), 12852-12860.

- Olins, P. O., Rangwala, S. H. (1989): A Novel Sequence Element Derived from Bacteriophage T7 mRNA Acts as an Enhancer of Translation of the lacZ Gene in Escherichia coli. *Journal of Biological Chemistry*, 264(29), 16973-16976.
- Oliveira, G. P., Andrade, A. C., Rodrigues, R. A., Arantes, T. S., Boratto, P. V., Silva, L. K., Dornas, F. P., Trindade, G. S., Drumond, B. P., La Scola, B., Kroon, E. G., Abrahão, J. S. (2017): Promoter Motifs in NCLDVs: An Evolutionary Perspective. *Viruses*, 9(1), 16.
- Pacha, R. F., Meis, R. J., Condit, R. C. (1990): Structure and Expression of the Vaccinia Virus Gene Which Prevents Virus-Induced Breakdown of RNA. *Journal of Virology*, 64(8), 3853-3863.
- Paoletti, E., Moss, B. (1974): Two Nucleic Acid-dependent Nucleoside Triphosphate Phosphohydrolases from Vaccinia Virus. Nucleotide substrate and polynucleotide cofactor specificities. *Journal of Biological Chemistry*, 249(10), 3281-3286.
- Paran, N., De Silva, F. S., Senkevich, T. G., Moss, B. (2009): Cellular DNA Ligase I Is Recruited to Cytoplasmic Vaccinia Virus Factories and Masks the Role of the Vaccinia Ligase in Viral DNA Replication. *Cell Host & Microbe*, 6(6), 563-569.
- Parrish, S., Moss, B. (2007): Characterization of a Second Vaccinia Virus mRNA-Decapping Enzyme Conserved in Poxviruses. *Journal of Virology*, 81(23), 12973-12978.
- Parrish, S., Resch, W., Moss, B. (2007): Vaccinia virus D10 protein has mRNA decapping activity, providing a mechanism for control of host and viral gene expression. *Proceedings of the National Academy of Sciences*, 104(7), 2139-2144.
- Parsons, B. L., Pickup, D. J. (1990): Transcription of Orthopoxvirus Telomeres at Late Times during Infection. *Virology*, 175(1), 69-80.
- Pascual-Ahuir, A., Proft, M. (2012): Quantification of Protein-DNA Interactions by In Vivo Chromatin Immunoprecipitation in Yeast. *Methods in Molecular Biology*, 809, 149-156.
- Passarelli, A. L., Kovacs, G. R., Moss, B. (1996): Transcription of a Vaccinia Virus Late Promoter Template: Requirement for the Product of the A2L Intermediate-Stage Gene. *Journal of Virology*, 70(7), 4444-4450.
- Patel, D. D., Pickup, D. J. (1989): The Second-Largest Subunit of the Poxvirus RNA Polymerase Is Similar to the Corresponding Subunits of Prokaryotic and Eucaryotic RNA Polymerases. *Journal of Virology*, 63(3), 1076-1086.
- Patel, G. P., Ma, S., Bag, J. (2005): The autoregulatory translational control element of poly(A)-binding protein mRNA forms a heteromeric ribonucleoprotein complex. *Nucleic Acids Research*, 33(22), 7074-7089.
- Pei, Y., Ho, C. K., Schwer, B., Shuman, S. (1999): Mutational Analyses of Yeast RNA Triphosphatases Highlight a Common Mechanism of Metal-dependent NTP Hydrolysis and a Means of Targeting Enzymes to Pre-mRNAs in Vivo by Fusion to the Guanylyltransferase Component of the Capping Apparatus. *Journal of Biological Chemistry*, 274(41), 28865-28874.
- Pejznochová, M. (2004): *Pravděpodobná cytoplasmatická belikasa lineárních plasmidů pGKL*. Diplomová práce, Univerzita Karlova, Praha.
- Philippe, N., Legendre, M., Doutre, G., Couté, Y., Poirot, O., Lescot, M., Arslan, D., Seltzer, V., Bertaux, L., Bruley, C., Garin, J., Claverie, J. M., Abergel, C. (2013): Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science*, 341(6143), 281-286.
- Piacente, S., Christen, L., Dickerman, B., Mohamed, M. R., Niles, E. G. (2008): Determinants of vaccinia virus early gene transcription termination. *Virology*, 376(1), 211-224.
- Piacente, S. C., Christen, L. A., Mohamed, M. R., Niles, E. G. (2003): Effect of selected mutations in the C-terminal region of the vaccinia virus nucleoside triphosphate phosphohydrolase I on binding to the H4L subunit of the viral RNA polymerase and early gene transcription termination in vitro. *Virology*, 310(1), 109-117.
- Pogo, B. G., O'Shea, M. T. (1978): The Mode of Replication of Vaccinia Virus DNA. *Virology*, 84(1), 1-8.
- Pogo, B. G. T., Berkowitz, E. M., Dales, S. (1984): Investigation of Vaccinia Virus DNA Replication Employing a Conditional Lethal Mutant Defective in DNA. *Virology*, 132(2), 436-444.
- Pospíšek, M., Palková, Z. (1991): Microisolation of yeast nucleic acids on the microtitre plate without using lytic enzymes. *Nucleic Acids Research*, 19(18), 5083-5083.
- Prescott, D. M., Kates, J., Kirkpatrick, J. B. (1971): Replication of Vaccinia Virus DNA in Enucleated L-cells. *Journal of Molecular Biology*, 59(3), 505-508.
- Priet, S., Lartigue, A., Debart, F., Claverie, J. M., Abergel, C. (2015): mRNA maturation in giant viruses: variation on a theme. *Nucleic Acids Research*, 43(7), 3776-3788.
- Prins, C., Cresawn, S. G., Condit, R. C. (2004): An Isatin- $\beta$ -thiosemicarbazone-resistant Vaccinia Virus Containing a Mutation in the Second Largest Subunit of the Viral RNA Polymerase Is Defective in Transcription Elongation. *Journal of Biological Chemistry*, 279(43), 44858-44871.
- Pritham, E. J., Putliwala, T., Feschotte, C. (2007): Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene*, 390(1-2), 3-17.
- Punjabi, A., Boyle, K., DeMasi, J., Grubisha, O., Unger, B., Khanna, M., Traktman, P. (2001): Clustered Charge-to-Alanine Mutagenesis of the Vaccinia Virus A20 Gene: Temperature-Sensitive Mutants Have a DNA-Minus Phenotype and Are Defective in the Production of Processive DNA Polymerase Activity. *Journal of Virology*, 75(24), 12308-12318.

- Raghavan, R., Kelkar, Y. D., Ochman, H. (2012): A selective force favoring increased G+C content in bacterial genes. *Proceedings of the National Academy of Sciences of the United States of America*, 109(36), 14504-14507.
- Rajyaguru, P., She, M., Parker, R. (2012): Scd6 targets eIF4G to repress translation: RGG-motif proteins as a class of eIF4G-binding proteins. *Molecular Cell*, 45(2), 244-254.
- Ray-Soni, A., Bellecourt, M. J., Landick, R. (2016): Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annual Review of Biochemistry*, 85, 319-347.
- Razzell, P. (1999): The origins of vaccinia virus - A brief comment. *Social History of Medicine*, 12(1), 141-141.
- Resch, W., Hixson, K. K., Moore, R. J., Lipton, M. S., Moss, B. (2007): Protein composition of the vaccinia virus mature virion. *Virology*, 358(1), 233-247.
- Reuter, J. S., Mathews, D. H. (2010): RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*, 11, 129.
- Rhee, H. S., Pugh, B. F. (2011): Comprehensive Genome-wide Protein-DNA Interactions Detected at Single-Nucleotide Resolution. *Cell*, 147(6), 1408-1419.
- Rodriguez, J. F., Kahn, J. S., Esteban, M. (1986): Molecular cloning, encoding sequence, and expression of vaccinia virus nucleic acid-dependent nucleoside triphosphatase gene. *Proceedings of the National Academy of Sciences of the United States of America*, 83(24), 9566-9570.
- Rodríguez, J. M., Salas, M. L. (2013): African swine fever virus transcription. *Virus Research*, 173(1), 15-28.
- Rohe, M., Schründer, J., Tudzynski, P., Meinhardt, F. (1992): Phylogenetic relationships of linear, protein-primed replicating genomes. *Current Genetics*, 21(2), 173-176.
- Rohrmann, G., Moss, B. (1985): Transcription of Vaccinia Virus Early Genes by a Template-Dependent Soluble Extract of Purified Virions. *Journal of Virology*, 56(2), 349-355.
- Rohrmann, G., Yuen, L., Moss, B. (1986): Transcription of Vaccinia Virus Early Genes by Enzymes Isolated from Vaccinia Virions Terminates Downstream of a Regulatory Sequence. *Cell*, 46(7), 1029-1035.
- Rochester, S. C., Traktman, P. (1998): Characterization of the Single-Stranded DNA Binding Protein Encoded by the Vaccinia Virus I3 Gene. *Journal of Virology*, 72(4), 2917-2926.
- Romanos, M. A., Boyd, A. (1988): A transcriptional barrier to expression of cloned toxin genes of the linear plasmid k1 of *Kluyveromyces lactis*: evidence that native k1 has novel promoters. *Nucleic Acids Research*, 16(15), 7333-7350.
- Rosales, R., Harris, N., Ahn, B. Y., Moss, B. (1994a): Purification and Identification of a Vaccinia Virus-encoded Intermediate Stage Promoter-specific Transcription Factor That Has Homology to Eukaryotic Transcription Factor SII (TFIIS) and an Additional Role as a Viral RNA Polymerase Subunit. *Journal of Biological Chemistry*, 269(19), 14260-14267.
- Rosales, R., Sutter, G., Moss, B. (1994b): A cellular factor is required for transcription of vaccinia viral intermediate-stage genes. *Proceedings of the National Academy of Sciences of the United States of America*, 91(9), 3794-3798.
- Rosel, J. L., Earl, P. L., Weir, J. P., Moss, B. (1986): Conserved TAAATG Sequence at the Transcriptional and Translational Initiation Sites of Vaccinia Virus Late Genes Deduced by Structural and Functional Analysis of the HindIII H Genome Fragment. *Journal of Virology*, 60(2), 436-449.
- Ruprich-Robert, G., Thuriaux, P. (2010): Non-canonical DNA transcription enzymes and the conservation of two-barrel RNA polymerases. *Nucleic Acids Research*, 38(14), 4559-4569.
- Saha, N., Shuman, S. (2001): Effects of Alanine Cluster Mutations in the D12 Subunit of Vaccinia Virus mRNA (Guanine-N7) Methyltransferase. *Virology*, 287(1), 40-48.
- Saha, N., Shuman, S., Schwer, B. (2003): Yeast-Based Genetic System for Functional Analysis of Poxvirus mRNA Cap Methyltransferase. *Journal of Virology*, 77(13), 7300-7307.
- Saha, N., Schwer, B., Shuman, S. (1999): Characterization of Human, *Schizosaccharomyces pombe*, and *Candida albicans* mRNA Cap Methyltransferases and Complete Replacement of the Yeast Capping Apparatus by Mammalian Enzymes. *Journal of Biological Chemistry*, 274(23), 16553-16562.
- Sachs, A. B., Davis, R. W. (1989): The Poly(A) Binding Protein Is Required for Poly(A) Shortening and 60S Ribosomal Subunit-Dependent Translation Initiation. *Cell*, 58(5), 857-867.
- Sachs, A. B., Davis, R. W., Kornberg, R. D. (1987): A Single Domain of Yeast Poly(A)-Binding Protein Is Necessary and Sufficient for RNA Binding and Cell Viability. *Molecular and Cellular Biology*, 7(9), 3268-3276.
- Salem, T. Z., Turney, C. M., Wang, L., Xue, J., Wan, X. F., Cheng, X. W. (2008): Transcriptional analysis of a major capsid protein gene from *Spodoptera exigua* ascovirus 5a. *Archives of Virology*, 153(1), 149-162.
- Sanz, P., Moss, B. (1998): A New Vaccinia Virus Intermediate Transcription Factor. *Journal of Virology*, 72(8), 6880-6883.
- Sanz, P., Moss, B. (1999): Identification of a transcription factor, encoded by two vaccinia virus early genes, that regulates the intermediate stage of viral gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 96(6), 2692-2697.
- Sarov, I., Joklik, W. K. (1973): Isolation and Characterization of Intermediates in Vaccinia Virus Morphogenesis. *Virology*, 52(1), 223-233.
- Satwika, D., Klassen, R., Meinhardt, F. (2012): Anticodon nuclease encoding virus-like elements in yeast. *Applied Microbiology and Biotechnology*, 96(2), 345-356.

- Segal, S. P., Dunckley, T., Parker, R. (2006): Sbp1p Affects Translational Repression and Decapping in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 26(13), 5120-5130.
- Sekyrová, P. (2004): *Proteom lineárních plasmidů kvasinky Kluyveromyces lactis*. Diplomová práce, Univerzita Karlova, Praha.
- Senkevich, T. G., Bruno, D., Martens, C., Porcella, S. F., Wolf, Y. I., Moss, B. (2015): Mapping vaccinia virus DNA replication origins at nucleotide level by deep sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, 112(35), 10908-10913.
- Senkevich, T. G., Katsafanas, G. C., Weisberg, A., Olano, L. R., Moss, B. (2017): Identification of Vaccinia Virus Replisome and Transcriptome Proteins by Isolation of Proteins on Nascent DNA Coupled with Mass Spectrometry. *Journal of Virology*, 91(19), e01015-17.
- Severinov, K., Mooney, R., Darst, S. A., Landick, R. (1997): Tethering of the Large Subunits of Escherichia coli RNA Polymerase. *Journal of Biological Chemistry*, 272(39), 24137-24140.
- Severinov, K., Mustaev, A., Kukarin, A., Muzzin, O., Bass, I., Darst, S. A., Goldfarb, A. (1996): Structural modules of the large subunits of RNA polymerase. Introducing archaeobacterial and chloroplast split sites in the beta and beta' subunits of Escherichia coli RNA polymerase. *Journal of Biological Chemistry*, 271(44), 27969-27974.
- Shatkin, A. J. (1976): Capping of Eucaryotic mRNAs. *Cell*, 9(4 Pt 2), 645-653.
- Shatzer, A. N., Kato, S. E., Condit, R. C. (2008): Phenotypic analysis of a temperature sensitive mutant in the large subunit of the vaccinia virus mRNA capping enzyme. *Virology*, 375(1), 236-252.
- Sheffield, F. W., Bauer, D. J., Stephenson, S. M. (1960): The Protection of Tissue Cultures by Isatin  $\beta$ -thiosemicarbazone from the Cytopathic Effects of Certain Pox Viruses. *British Journal of Experimental Pathology*, 41(6), 638-647.
- Shevchenko, A., Tomas, H., Havliš, J., Olsen, J. V., Mann, M. (2006): In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, 1(6), 2856-2860.
- Shinshi, H., Miwa, M., Kato, K., Noguchi, M., Matsushima, T., Sugimura, T. (1976a): A Novel Phosphodiesterase from Cultured Tobacco Cells. *Biochemistry*, 15(10), 2185-2190.
- Shinshi, H., Miwa, M., Sugimura, T. (1976b): Enzyme cleaving the 5'-terminal methylated blocked structure of messenger RNA. *FEBS Letters*, 65(2), 254-257.
- Shirokikh, N. E., Spirin, A. S. (2008): Poly(A) leader of eukaryotic mRNA bypasses the dependence of translation on initiation factors. *Proceedings of the National Academy of Sciences of the United States of America*, 105(31), 10738-10743.
- Shuman, S. (1989): Functional Domains of Vaccinia Virus mRNA Capping Enzyme. Analysis by limited tryptic digestion. *Journal of Biological Chemistry*, 264(16), 9690-9695.
- Shuman, S. (1992): Vaccinia virus RNA helicase: An essential enzyme related to the DE-H family of RNA-dependent NTPases. *Proceedings of the National Academy of Sciences of the United States of America*, 89(22), 10935-10939.
- Shuman, S. (1995): Capping enzyme in eukaryotic mRNA synthesis. *Progress in Nucleic Acid Research and Molecular Biology*, 50, 101-129.
- Shuman, S. (2002): What messenger RNA capping tells us about eukaryotic evolution. *Nature Reviews Molecular Cell Biology*, 3(8), 619-625.
- Shuman, S., Broyles, S. S., Moss, B. (1987): Purification and Characterization of a Transcription Termination Factor from Vaccinia Virions. *Journal of Biological Chemistry*, 262(25), 12372-12380.
- Shuman, S., Morham, S. G. (1990): Domain structure of vaccinia virus mRNA capping enzyme. Activity of the Mr 95,000 subunit expressed in Escherichia coli. *Journal of Biological Chemistry*, 265(20), 11967-11972.
- Shuman, S., Moss, B. (1988): Factor-dependent Transcription Termination by Vaccinia Virus RNA Polymerase - Evidence That the cis-Acting Termination Signal Is in Nascent RNA. *Journal of Biological Chemistry*, 263(13), 6220-6225.
- Shuman, S., Moss, B. (1989): Bromouridine Triphosphate Inhibits Transcription Termination and mRNA Release by Vaccinia Virions. *Journal of Biological Chemistry*, 264(35), 21356-21360.
- Shuman, S., Schwer, B. (1995): RNA capping enzyme and DNA ligase: a superfamily of covalent nucleotidyl transferases. *Molecular Microbiology*, 17(3), 405-410.
- Shuman, S., Surks, M., Furneaux, H., Hurwitz, J. (1980): Purification and Characterization of a GTP-Pyrophosphate Exchange Activity from Vaccinia Virions. Association of the GTP-pyrophosphate exchange activity with vaccinia mRNA guanylyltransferase. RNA (guanine-7-)methyltransferase complex (capping enzyme). *Journal of Biological Chemistry*, 255(23), 11588-11598.
- Schaffrath, R., Meacock, P. A. (1995): Kluyveromyces lactis Killer Plasmid pGKL2: Molecular Analysis of an Essential Gene, ORF5. *Yeast*, 11(7), 615-628.
- Schaffrath, R., Meacock, P. A. (2001): An SSB encoded by and operating on linear killer plasmids from Kluyveromyces lactis. *Yeast*, 18(13), 1239-1247.
- Schaffrath, R., Meinhardt, F., Meacock, P. A. (1997): ORF7 of yeast plasmid pGKL2: analysis of gene expression in vivo. *Current Genetics*, 31(2), 190-192.
- Schaffrath, R., Soond, S. M., Meacock, P. A. (1995a): Cytoplasmic gene expression in yeast: a plasmid-encoded transcription system in Kluyveromyces lactis. *Biochemical Society Transactions*, 23(1), 128S.

- Schaffrath, R., Soond, S. M., Meacock, P. A. (1995b): The DNA and RNA polymerase genes of yeast plasmid pGKL2 are essential loci for plasmid integrity and maintenance. *Microbiology*, 141 (Pt 10), 2591-2599.
- Schaffrath, R., Stark, M. J., Gunge, N., Meinhardt, F. (1992): Kluyveromyces lactis killer system: ORF1 of pGKL2 has no function in immunity expression and is dispensable for killer plasmid replication and maintenance. *Current Genetics*, 21(4-5), 357-363.
- Schickel, J., Helmig, C., Meinhardt, F. (1996): Kluyveromyces lactis killer system: analysis of cytoplasmic promoters of the linear plasmids. *Nucleic Acids Research*, 24(10), 1879-1886.
- Schmidt, F. I., Bleck, C. K. E., Reh, L., Novy, K., Wollscheid, B., Helenius, A., Stahlberg, H., Mercer, J. (2013): Vaccinia Virus Entry Is Followed by Core Activation and Proteasome-Mediated Release of the Immunomodulatory Effector VH1 from Lateral Bodies. *Cell Reports*, 4(3), 464-476.
- Schmidt, W. M., Mueller, M. W. (1999): CapSelect: a highly sensitive method for 5' CAP-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nucleic Acids Research*, 27(21), e31.
- Schnierle, B. S., Gershon, P. D., Moss, B. (1992): Cap-specific mRNA (nucleoside-O2'-)-methyltransferase and poly(A) polymerase stimulatory activities of vaccinia virus are mediated by a single protein. *Proceedings of the National Academy of Sciences of the United States of America*, 89(7), 2897-2901.
- Schnierle, B. S., Moss, B. (1992): Vaccinia Virus-Mediated Inhibition of Host Protein Synthesis Involves Neither Degradation nor Underphosphorylation of Components of the Cap-Binding Eukaryotic Translation Initiation Factor Complex eIF-4F. *Virology*, 188(2), 931-933.
- Schründer, J., Gunge, N., Meinhardt, F. (1996): Extranuclear expression of the bacterial xylose isomerase (xylA) and the UDP-glucose dehydrogenase (hasB) genes in yeast with Kluyveromyces lactis linear killer plasmids as vectors. *Current Microbiology*, 33(5), 323-330.
- Schründer, J., Meinhardt, F. (1995): An extranuclear expression system for analysis of cytoplasmic promoters of yeast linear killer plasmids. *Plasmid*, 33(2), 139-151.
- Schuch, R., Agelopoulos, K., Neumann, A., Brandt, B., Bürger, H., Korsching, E. (2012): Site-specific chromatin immunoprecipitation: a selective method to individually analyze neighboring transcription factor binding sites in vivo. *BMC Research Notes*, 5, 109.
- Schwer, B., Hausmann, S., Schneider, S., Shuman, S. (2006): Poxvirus mRNA cap methyltransferase. Bypass of the requirement for the stimulatory subunit by mutations in the catalytic subunit and evidence for intersubunit allostery. *Journal of Biological Chemistry*, 281(28), 18953-18960.
- Schwer, B., Saha, N., Mao, X., Chen, H. W., Shuman, S. (2000): Structure-Function Analysis of Yeast mRNA Cap Methyltransferase and High-Copy Suppression of Conditional Mutants by AdoMet Synthase and the Ubiquitin Conjugating Enzyme Cdc34p. *Genetics*, 155(4), 1561-1576.
- Schwer, B., Stunnenberg, H. G. (1988): Vaccinia virus late transcripts generated in vitro have a poly(A) head. *The EMBO Journal*, 7(4), 1183-1190.
- Schwer, B., Visca, P., Vos, J. C., Stunnenberg, H. G. (1987): Discontinuous Transcription or RNA Processing of Vaccinia Virus Late Messengers Results in a 5' Poly(A) Leader. *Cell*, 50(2), 163-169.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., Higgins, D. G. (2011): Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, 539.
- Simpson, D. A., Condit, R. C. (1995): Vaccinia Virus Gene A18R Encodes an Essential DNA Helicase. *Journal of Virology*, 69(10), 6131-6139.
- Sivan, G., Weisberg, A. S., Americo, J. L., Moss, B. (2016): Retrograde Transport from Early Endosomes to the trans-Golgi Network Enables Membrane Wrapping and Egress of Vaccinia Virus Virions. *Journal of Virology*, 90(19), 8891-8905.
- Smith, G. L., Law, M. (2004): The exit of Vaccinia virus from infected cells. *Virus Research*, 106(2), 189-197.
- Smith, G. L., Vanderplassen, A., Law, M. (2002): The formation and function of extracellular enveloped vaccinia virus. *Journal of General Virology*, 83, 2915-2931.
- Smits, W. K., Grossman, A. D. (2010): The Transcriptional Regulator Rok Binds A+T-Rich DNA and Is Involved in Repression of a Mobile Genetic Element in Bacillus subtilis. *PLoS Genetics*, 6(11), e1001207.
- Sonntag, K. C., Darai, G. (1995): Evolution of Viral DNA-Dependent RNA Polymerases. *Virus Genes*, 11(2-3), 271-284.
- Sor, F., Fukuhara, H. (1985): Structure of a linear plasmid of the yeast Kluyveromyces lactis; Compact organization of the killer genome. *Current Genetics*, 9(2), 147-155.
- Spencer, E., Shuman, S., Hurwitz, J. (1980): Purification and Properties of Vaccinia Virus DNA-dependent RNA Polymerase. *Journal of Biological Chemistry*, 255(11), 5388-5395.
- Sridhar, P., Condit, R. C. (1983): Selection for Temperature-Sensitive Mutations in Specific Vaccinia Virus Genes: Isolation and Characterization of a Virus Mutant Which Encodes a Phosphonoacetic Acid-Resistant, Temperature-Sensitive DNA Polymerase. *Virology*, 128(2), 444-457.
- Stam, J. C., Kwakman, J., Meijer, M., Stuitje, A. R. (1986): Efficient isolation of the linear DNA killer plasmid of Kluyveromyces lactis: evidence for location and expression in the cytoplasm and characterization of their terminally bound proteins. *Nucleic Acids Research*, 14(17), 6871-6884.
- Stark, M. J., Boyd, A., Mileham, A. J., Romanos, M. A. (1990): The plasmid-encoded killer system of Kluyveromyces lactis: a review. *Yeast*, 6(1), 1-29.



- Stark, M. J. R., Boyd, A.** (1986): The killer toxin of *Kluyveromyces fragilis*: characterization of the toxin subunits and identification of the genes which encode them. *The EMBO Journal*, 5(8), 1995-2002.
- Su, M.-J., Bablanian, R.** (1990): Polyadenylated RNA Sequences from Vaccinia Virus-Infected Cells Selectively Inhibit Translation in a Cell-Free System: Structural Properties and Mechanism of Inhibition. *Virology*, 179(2), 679-693.
- Suhre, K., Audic, S., Claverie, J. M.** (2005): Mimivirus gene promoters exhibit an unprecedented conservation among all eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, 102(41), 14689-14693.
- Sýkora, M.** (2013): *Charakterizace transkripčního aparátu lineárních plasmidů kvasinky Kluyveromyces fragilis*. Diplomová práce, Univerzita Karlova, Praha.
- Szajner, P., Weisberg, A. S., Lebowitz, J., Heuser, J., Moss, B.** (2005): External scaffold of spherical immature poxvirus particles is made of protein trimers, forming a honeycomb lattice. *The Journal of Cell Biology*, 170(6), 971-981.
- Takeda, M., Hiraishi, H., Takesako, T., Tanase, S., Gunge, N.** (1996): The terminal protein of the linear DNA plasmid pGKL2 shares an N-terminal domain of the plasmid-encoded DNA polymerase. *Yeast*, 12(3), 241-246.
- Tanner, N. K., Cordin, O., Banroques, J., Doère, M., Linder, P.** (2003): The Q Motif: A Newly Identified Motif in DEAD Box Helicases May Regulate ATP Binding and Hydrolysis. *Molecular Cell*, 11(1), 127-138.
- Tarun, S. Z., Jr., Sachs, A. B.** (1996): Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *The EMBO Journal*, 15(24), 7168-7177.
- Tate, J., Boldt, R. L., McFadden, B. D., D'Costa, S. M., Lewandowski, N. M., Shatzer, A. N., Gollnick, P., Condit, R. C.** (2016): Biochemical analysis of the multifunctional vaccinia mRNA capping enzyme encoded by a temperature sensitive virus mutant. *Virology*, 487, 27-40.
- Tate, J., Gollnick, P.** (2011): Role of Forward Translocation in Nucleoside Triphosphate Phosphohydrolase I (NPH I)-mediated Transcription Termination of Vaccinia Virus Early Genes. *Journal of Biological Chemistry*, 286(52), 44764-44775.
- Tate, J., Gollnick, P.** (2015): The role of vaccinia termination factor and cis-acting elements in vaccinia virus early gene transcription termination. *Virology*, 485, 179-188.
- Taylor, S. D., Solem, A., Kawaoka, J., Pyle, A. M.** (2010): The NPH-II Helicase Displays Efficient DNA·RNA Helicase Activity and a Pronounced Purine Sequence Bias. *The Journal of Biological Chemistry*, 285(15), 11692-11703.
- Tiggemann, M., Jeske, S., Larsen, M., Meinhardt, F.** (2001): *Kluyveromyces fragilis* cytoplasmic plasmid pGKL2: heterologous expression of Orf3p and proof of guanylyltransferase and mRNA-triphosphatase activities. *Yeast*, 18(9), 815-825.
- Tokunaga, M., Kawamura, A., Hishinuma, F.** (1989): Expression of pGKL killer 28K subunit in *Saccharomyces cerevisiae*: identification of 28K subunit as a killer protein. *Nucleic Acids Research*, 17(9), 3435-3446.
- Tokunaga, M., Wada, N., Hishinuma, F.** (1987): Expression and identification of immunity determinants on linear DNA killer plasmids pGKL1 and pGKL2 in *Kluyveromyces fragilis*. *Nucleic Acids Research*, 15(3), 1031-1046.
- Tolonen, N., Doglio, L., Schleich, S., Krijnse Locker, J.** (2001): Vaccinia Virus DNA Replication Occurs in Endoplasmic Reticulum-enclosed Cytoplasmic Mini-Nuclei. *Molecular Biology of the Cell*, 12(7), 2031-2046.
- Tommasino, M., Ricci, S., Galeotti, C. L.** (1988): Genome organization of the killer plasmid pGKL2 from *Kluyveromyces fragilis*. *Nucleic Acids Research*, 16(13), 5863-5878.
- Toulokhonov, I., Landick, R.** (2006): The Role of the Lid Element in Transcription by *E. coli* RNA Polymerase. *Journal of Molecular Biology*, 361(4), 644-658.
- Traktman, P., Caligiuri, A., Jesty, S. A., Liu, K., Sankar, U.** (1995): Temperature-Sensitive Mutants with Lesions in the Vaccinia Virus F10 Kinase Undergo Arrest at the Earliest Stage of Virion Morphogenesis. *Journal of Virology*, 69(10), 6581-6587.
- Tutas, D. J., Paoletti, E.** (1977): Purification and Characterization of Core-associated Polynucleotide 5'-Triphosphatase from Vaccinia Virus. *Journal of Biological Chemistry*, 252(9), 3092-3098.
- Unger, B., Mercer, J., Boyle, K. A., Traktman, P.** (2013): Biogenesis of the Vaccinia Virus Membrane: Genetic and Ultrastructural Analysis of the Contributions of the A14 and A17 Proteins. *Journal of Virology*, 87(2), 1083-1097.
- Unniraman, S., Prakash, R., Nagaraja, V.** (2001): Alternate Paradigm for Intrinsic Transcription Termination in Eubacteria. *Journal of Biological Chemistry*, 276(45), 41850-41855.
- Upton, C., Slack, S., Hunter, A. L., Ehlers, A., Roper, R. L.** (2003): Poxvirus Orthologous Clusters: toward Defining the Minimum Essential Poxvirus Genome. *Journal of Virology*, 77(13), 7590-7600.
- Urushibara, T., Furuichi, Y., Nishimura, C., Miura, K.** (1975): A MODIFIED STRUCTURE AT THE 5'-TERMINUS OF mRNA OF VACCINIA VIRUS. *FEBS Letters*, 49(3), 385-389.
- Van Dyck, E., Foury, F., Stillman, B., Brill, S. J.** (1992): A single-stranded DNA binding protein required for mitochondrial DNA replication in *S. cerevisiae* is homologous to *E. coli* SSB. *The EMBO Journal*, 11(9), 3421-3430.

- Vanslyke, J. K., Lee, P., Wilson, E. M., Hruby, D. E. (1993): Isolation and Analysis of Vaccinia Virus Previrions. *Virus Genes*, 7(4), 311-324.
- Vassef, A., Mars, M., Dru, A., Plucienniczak, A., Streeck, R. E., Beaud, G. (1985): Isolation of cis-Acting Vaccinia Virus DNA Fragments Promoting the Expression of Herpes Simplex Virus Thymidine Kinase by Recombinant Viruses. *Journal of Virology*, 55(1), 163-172.
- Vassilyev, D. G., Vassilyeva, M. N., Perederina, A., Tahirov, T. H., Artsimovitch, I. (2007): Structural basis for transcription elongation by bacterial RNA polymerase. *Nature*, 448(7150), 157-162.
- Venkatesan, S., Gershowitz, A., Moss, B. (1980): Modification of the 5' End of mRNA. Association of RNA triphosphatase with the RNA guanylyltransferase-RNA (guanine-7-)-methyltransferase complex from vaccinia virus. *Journal of Biological Chemistry*, 255(3), 903-908.
- Venkatesan, S., Moss, B. (1981): In Vitro Transcription of the Inverted Terminal Repetition of the Vaccinia Virus Genome: Correspondence of Initiation and Cap Sites. *Journal of Virology*, 37(2), 738-747.
- Vopálenský, V. (1999): *Lineární plasmidy kvasinek*. Diplomová práce, Univerzita Karlova, Praha.
- Vopálenský, V. (2007): *Role 5' koncových struktur u eukaryotických mRNA*. Dizertační práce, Univerzita Karlova, Praha.
- Vopálenský, V., Sýkora, M., Mašek, T., Pospíšek, M. (2018): Messenger RNAs transcribed from yeast linear cytoplasmic plasmids possess unconventional 5' and 3' UTRs and suggest a novel mechanism of translation. *bioRxiv*, doi: 10.1101/325316.
- Vos, J. C., Saker, M., Stunnenberg, H. G. (1991a): Promoter Melting by a Stage-Specific Vaccinia Virus Transcription Factor Is Independent of the Presence of RNA Polymerase. *Cell*, 65(1), 105-113.
- Vos, J. C., Saker, M., Stunnenberg, H. G. (1991b): Vaccinia virus capping enzyme is a transcription initiation factor. *The EMBO Journal*, 10(9), 2553-2558.
- Vos, J. C., Stunnenberg, H. G. (1988): Derepression of a novel class of vaccinia virus genes upon DNA replication. *The EMBO Journal*, 7(11), 3487-3492.
- Walsh, D. (2017): Poxviruses: Slipping and sliding through transcription and translation. *PLoS Pathogens*, 13(11), e1006634.
- Walsh, D., Arias, C., Perez, C., Halladin, D., Escandon, M., Ueda, T., Watanabe-Fukunaga, R., Fukunaga, R., Mohr, I. (2008): Eukaryotic translation initiation factor 4F architectural alterations accompany translation initiation factor redistribution in poxvirus-infected cells. *Molecular and Cellular Biology*, 28(8), 2648-2658.
- Wang, M. (2011): Seqool – A sequence analysis tool for signal search, pattern recognition and sequence statistics. Verze 3.1, <http://biossc.de/seqool>.
- Wang, S. P., Shuman, S. (1997): Structure-Function Analysis of the mRNA Cap Methyltransferase of *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 272(23), 14683-14689.
- Wei, C. M., Moss, B. (1974): Methylation of Newly Synthesized Viral Messenger RNA by an Enzyme in Vaccinia Virus. *Proceedings of the National Academy of Sciences of the United States of America*, 71(8), 3014-3018.
- Wei, C. M., Moss, B. (1975): Methylated Nucleotides Block 5'-Terminus of Vaccinia Virus Messenger RNA. *Proceedings of the National Academy of Sciences of the United States of America*, 72(1), 318-322.
- Weir, J. P., Moss, B. (1983): Nucleotide Sequence of the Vaccinia Virus Thymidine Kinase Gene and the Nature of Spontaneous Frameshift Mutations. *Journal of Virology*, 46(2), 530-537.
- Weir, J. P., Moss, B. (1984): Regulation of Expression and Nucleotide Sequence of a Late Vaccinia Virus Gene. *Journal of Virology*, 51(3), 662-669.
- Weir, J. P., Moss, B. (1987): Determination of the Transcriptional Regulatory Region of a Vaccinia Virus Late Gene. *Journal of Virology*, 61(1), 75-80.
- Welnowska, E., Castelló, A., Moral, P., Carrasco, L. (2009): Translation of mRNAs from Vesicular Stomatitis Virus and Vaccinia Virus Is Differentially Blocked in Cells with Depletion of eIF4GI and/or eIF4GII. *Journal of Molecular Biology*, 394(3), 506-521.
- Werner, F., Grohmann, D. (2011): Evolution of multisubunit RNA polymerases in the three domains of life. *Nature Reviews Microbiology*, 9(2), 85-98.
- Wickner, R. B., Fujimura, T., Esteban, R. (2013): Viruses and prions of *Saccharomyces cerevisiae*. *Advances in Virus Research*, 86, 1-36.
- Wilson, D. W., Meacock, P. A. (1988): Extranuclear gene expression in yeast: evidence for a plasmid-encoded RNA polymerase of unique structure. *Nucleic Acids Research*, 16(16), 8097-8112.
- Wittek, R., Cooper, J. A., Barbosa, E., Moss, B. (1980): Expression of the Vaccinia Virus Genome: Analysis and Mapping of mRNAs Encoded within the Inverted Terminal Repetition. *Cell*, 21(2), 487-493.
- Wittek, R., Moss, B. (1980): Tandem Repeats within the Inverted Terminal Repetition of Vaccinia Virus DNA. *Cell*, 21(1), 277-284.
- Woods, D. R., Bevan, E. A. (1968): Studies on the Nature of the Killer Factor Produced by *Saccharomyces cerevisiae*. *Journal of General Microbiology*, 51(1), 115-126.
- Wright, C. F., Coroneos, A. M. (1993): Purification of the Late Transcription System of Vaccinia Virus: Identification of a Novel Transcription Factor. *Journal of Virology*, 67(12), 7264-7270.
- Wright, C. F., Coroneos, A. M. (1995): The H4 Subunit of Vaccinia Virus RNA Polymerase Is Not Required for Transcription Initiation at a Viral Late Promoter. *Journal of Virology*, 69(4), 2602-2604.

- Wright, C. F., Hubbs, A. E., Gunasinghe, S. K., Oswald, B. W. (1998): A Vaccinia Virus Late Transcription Factor Copurifies with a Factor That Binds to a Viral Late Promoter and Is Complemented by Extracts from Uninfected HeLa Cells. *Journal of Virology*, 72(2), 1446-1451.
- Wright, C. F., Keck, J. G., Tsai, M. M., Moss, B. (1991): A Transcription Factor for Expression of Vaccinia Virus Late Genes Is Encoded by an Intermediate Gene. *Journal of Virology*, 65(7), 3715-3720.
- Wright, C. F., Moss, B. (1987): In vitro synthesis of vaccinia virus late mRNA containing a 5' poly(A) leader sequence. *Proceedings of the National Academy of Sciences of the United States of America*, 84(24), 8883-8887.
- Wright, C. F., Moss, B. (1989): Identification of Factors Specific for Transcription of the Late Class of Vaccinia Virus Genes. *Journal of Virology*, 63(10), 4224-4233.
- Wright, C. F., Oswald, B. W., Dellis, S. (2001): Vaccinia Virus Late Transcription Is Activated in Vitro by Cellular Heterogeneous Nuclear Ribonucleoproteins. *Journal of Biological Chemistry*, 276(44), 40680-40686.
- Xia, X., MacKay, V., Yao, X., Wu, J., Miura, F., Ito, T., Morris, D. R. (2011): Translation initiation: a regulatory role for poly(A) tracts in front of the AUG codon in *Saccharomyces cerevisiae*. *Genetics*, 189(2), 469-478.
- Xiang, Y., Latner, D. R., Niles, E. G., Condit, R. C. (2000): Transcription Elongation Activity of the Vaccinia Virus J3 Protein in Vivo Is Independent of Poly(A) Polymerase Stimulation. *Virology*, 269(2), 356-369.
- Xiang, Y., Simpson, D. A., Spiegel, J., Zhou, A. M., Silverman, R. H., Condit, R. C. (1998): The Vaccinia Virus A18R DNA Helicase Is a Postreplicative Negative Transcription Elongation Factor. *Journal of Virology*, 72(9), 7012-7023.
- Yáñez, R. J., Rodríguez, J. M., Bournsnel, M., Rodríguez, J. F., Viñuela, E. (1993): Two putative African swine fever virus helicases similar to yeast 'DEAH' pre-mRNA processing proteins and vaccinia virus ATPases D11L and D6R. *Gene*, 134(2), 161-174.
- Yang, Z. (2018): Selective translation conferred by the poly(A) leader of vaccinia virus mRNAs. City: Manhattan, Kansas. National Institute of Health: Kansas State University, United States of America.
- Yang, Z., Bruno, D. P., Martens, C. A., Porcella, S. F., Moss, B. (2010): Simultaneous high-resolution analysis of vaccinia virus and host cell transcriptomes by deep RNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, 107(25), 11513-11518.
- Yang, Z., Bruno, D. P., Martens, C. A., Porcella, S. F., Moss, B. (2011a): Genome-Wide Analysis of the 5' and 3' Ends of Vaccinia Virus Early mRNAs Delineates Regulatory Sequences of Annotated and Anomalous Transcripts. *Journal of Virology*, 85(12), 5897-5909.
- Yang, Z., Cao, S., Martens, C. A., Porcella, S. F., Xie, Z., Ma, M., Shen, B., Moss, B. (2015): Deciphering Poxvirus Gene Expression by RNA Sequencing and Ribosome Profiling. *Journal of Virology*, 89(13), 6874-6886.
- Yang, Z., Martens, C. A., Bruno, D. P., Porcella, S. F., Moss, B. (2012): Pervasive Initiation and 3'-End Formation of Poxvirus Postreplicative RNAs. *Journal of Biological Chemistry*, 287(37), 31050-31060.
- Yang, Z., Reynolds, S. E., Martens, C. A., Bruno, D. P., Porcella, S. F., Moss, B. (2011b): Expression Profiling of the Intermediate and Late Stages of Poxvirus Replication. *Journal of Virology*, 85(19), 9899-9908.
- Yang, Z. L., Maruri-Avidal, L., Sisler, J., Stuart, C. A., Moss, B. (2013): Cascade regulation of vaccinia virus gene expression is modulated by multistage promoters. *Virology*, 447(1-2), 213-220.
- Yang, Z. L., Moss, B. (2009): Interaction of the Vaccinia Virus RNA Polymerase-Associated 94-Kilodalton Protein with the Early Transcription Factor. *Journal of Virology*, 83(23), 12018-12026.
- Yoder, J. D., Chen, T. S., Gagnier, C. R., Vemulapalli, S., Maier, C. S., Hruby, D. E. (2006): Pox proteomics: mass spectrometry analysis and identification of Vaccinia virion proteins. *Virology Journal*, 3, 10.
- Yu, L., Martins, A., Deng, L., Shuman, S. (1997): Structure-Function Analysis of the Triphosphatase Component of Vaccinia Virus mRNA Capping Enzyme. *Journal of Virology*, 71(12), 9837-9843.
- Yu, L., Shuman, S. (1996): Mutational Analysis of the RNA Triphosphatase Component of Vaccinia Virus mRNA Capping Enzyme. *Journal of Virology*, 70(9), 6162-6168.
- Yuen, L., Davison, A. J., Moss, B. (1987): Early promoter-binding factor from vaccinia virions. *Proceedings of the National Academy of Sciences of the United States of America*, 84(17), 6069-6073.
- Yuen, L., Moss, B. (1986): Multiple 3' Ends of mRNA Encoding Vaccinia Virus Growth Factor Occur within a Series of Repeated Sequences Downstream of T Clusters. *Journal of Virology*, 60(1), 320-323.
- Yuen, L., Moss, B. (1987): Oligonucleotide sequence signaling transcriptional termination of vaccinia virus early genes. *Proceedings of the National Academy of Sciences of the United States of America*, 84(18), 6417-6421.
- Yutin, N., Koonin, E. V. (2012): Hidden evolutionary complexity of Nucleo-Cytoplasmic Large DNA viruses of eukaryotes. *Virology Journal*, 9, 161.
- Yutin, N., Raoult, D., Koonin, E. V. (2013): Virophages, polintons, and transpovirons: a complex evolutionary network of diverse selfish genetic elements with different reproduction strategies. *Virology Journal*, 10, 158.
- Yutin, N., Shevchenko, S., Kapitonov, V., Krupović, M., Koonin, E. V. (2015): A novel group of diverse Polinton-like viruses discovered by metagenome analysis. *BMC Biology*, 13, 95.
- Zaborowska, I., Kellner, K., Henry, M., Meleady, P., Walsh, D. (2012): Recruitment of host translation initiation factor eIF4G by the Vaccinia Virus ssDNA-binding protein I3. *Virology*, 425(1), 11-22.
- Zaborowska, I., Walsh, D. (2009): PI3K Signaling Regulates Rapamycin-Insensitive Translation Initiation Complex Formation in Vaccinia Virus-Infected Cells. *Journal of Virology*, 83(8), 3988-3992.

- Zaghloul, H. A. H., Hice, R., Arensbarger, P., Federici, B. A.** (2017): Transcriptome Analysis of the *Spodoptera frugiperda* Ascovirus In Vivo Provides Insights into How Its Apoptosis Inhibitors and Caspase Promote Increased Synthesis of Viral Vesicles and Virion Progeny. *Journal of Virology*, 91(23), e00874-17.
- Zhang, Y.** (2008): I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, 9, 40.
- Zhang, Y., Keck, J. G., Moss, B.** (1992): Transcription of Viral Late Genes Is Dependent on Expression of the Viral Intermediate Gene G8R in Cells Infected with an Inducible Conditional-Lethal Mutant Vaccinia Virus. *Journal of Virology*, 66(11), 6470-6479.
- Zheng, S., Shuman, S.** (2008a): Mutational analysis of vaccinia virus mRNA cap (guanine-N7) methyltransferase reveals essential contributions of the N-terminal peptide that closes over the active site. *RNA*, 14(11), 2297-2304.
- Zheng, S., Shuman, S.** (2008b): Structure–function analysis of vaccinia virus mRNA cap (guanine-N7) methyltransferase. *RNA*, 14(4), 696-705.
- Zhu, M., Moore, T., Broyles, S. S.** (1998): A Cellular Protein Binds Vaccinia Virus Late Promoters and Activates Transcription In Vitro. *Journal of Virology*, 72(5), 3893-3899.

## 10 VYBRANÉ PUBLIKACE

### 10.1 KVASINKOVÉ CYTOPLASMATICKÉ LINEÁRNÍ PLASMIDY

#### 10.1.1 Transcription apparatus of the yeast virus-like elements: Architecture, function, and evolutionary origin

Michal Sýkora, Martin Pospíšek, Josef Novák, Silvia Mrvová, Libor Krásný a Václav Vopálenský

*PLoS Pathogens*, 22. října 2018

PMID: 30346988

DOI: <https://doi.org/10.1371/journal.ppat.1007377>

IF<sub>2018/2019</sub>: 6,463

Doplňkové informace: <https://ndownloader.figshare.com/articles/7237412/versions/1>

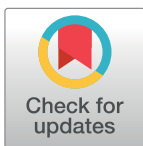
RESEARCH ARTICLE

# Transcription apparatus of the yeast virus-like elements: Architecture, function, and evolutionary origin

Michal Sýkora<sup>1</sup>, Martin Pospíšek<sup>1\*</sup>, Josef Novák<sup>1</sup>, Silvia Mrvová<sup>1</sup>, Libor Krásný<sup>2</sup>, Václav Vopálenský<sup>1\*</sup>

**1** Department of Genetics and Microbiology, Faculty of Science, Charles University, Prague, Czech Republic, **2** Institute of Microbiology, Academy of Sciences of the Czech Republic, Prague, Czech Republic

\* [martin@natur.cuni.cz](mailto:martin@natur.cuni.cz) (MP); [vasek@natur.cuni.cz](mailto:vasek@natur.cuni.cz) (VV)



## OPEN ACCESS

**Citation:** Sýkora M, Pospíšek M, Novák J, Mrvová S, Krásný L, Vopálenský V (2018) Transcription apparatus of the yeast virus-like elements: Architecture, function, and evolutionary origin. *PLoS Pathog* 14(10): e1007377. <https://doi.org/10.1371/journal.ppat.1007377>

**Editor:** Robert Belshaw, Plymouth University, UNITED KINGDOM

**Received:** May 21, 2018

**Accepted:** October 3, 2018

**Published:** October 22, 2018

**Copyright:** © 2018 Sýkora et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by Czech Science Foundation (P305/12/G034), Charles University Institutional Project (SVV-260426), and ELIXIR CZ Research Infrastructure Project (MEYS Grant No: LM2015047). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Extrachromosomal hereditary elements such as organelles, viruses, and plasmids are important for the cell fitness and survival. Their transcription is dependent on host cellular RNA polymerase (RNAP) or intrinsic RNAP encoded by these elements. The yeast *Kluyveromyces lactis* contains linear cytoplasmic DNA virus-like elements (VLEs, also known as linear plasmids) that bear genes encoding putative non-canonical two-subunit RNAP. Here, we describe the architecture and identify the evolutionary origin of this transcription machinery. We show that the two RNAP subunits interact *in vivo*, and this complex interacts with another two VLE-encoded proteins, namely the mRNA capping enzyme and a putative helicase. RNAP, mRNA capping enzyme and the helicase also interact with VLE-specific DNA *in vivo*. Further, we identify a promoter sequence element that causes 5' mRNA polyadenylation of VLE-specific transcripts *via* RNAP slippage at the transcription initiation site, and structural elements that precede the termination sites. As a result, we present a first model of the yeast virus-like element transcription initiation and intrinsic termination. Finally, we demonstrate that VLE RNAP and its promoters display high similarity to poxviral RNAP and promoters of early poxviral genes, respectively, thereby pointing to their evolutionary origin.

## Author summary

Yeast cytoplasmic double-stranded DNA virus-like elements (VLEs, also known as linear plasmids) were widely investigated in the past but the topic was almost entirely abandoned, partly due to an inability to express VLE-encoded proteins using conventional expression systems. In this study, we re-opened investigation of these elements focused on considerably underexplored nucleus-independent transcription of *K. lactis* VLEs. Using systematic *in vivo* study, we were able to characterize composition of the previously hypothesized VLE-encoded transcription complex. Further, we identified new DNA and RNA elements that were directly connected with formation of VLE-specific mRNA ends, as demonstrated by mutagenesis of these elements and its effect on VLE transcription *in vivo*. Finally, our phylogenetic and sequence analysis of VLE-encoded non-canonical



**Competing interests:** The authors have declared that no competing interests exist.

RNA polymerase and its promoters, respectively, suggested evolutionary relationship of VLE transcription machinery to much better explored transcription machinery of cytoplasmic double-stranded DNA viruses of the *Poxviridae* family.

## Introduction

Linear double-stranded DNA virus-like elements (VLEs) were found in the cytoplasm of several yeast species. Structural organization of these elements is quite uniform and they are often present as two or three differently sized DNA plasmids in yeast host cells [1]. Characteristic features of VLEs are terminal proteins covalently linked to the 5' ends of their DNA, terminal inverted repeats, and their cytoplasmic localization [2–4]. Yeast linear plasmids of *Kluyveromyces lactis*, termed pGKL1 (or K1) and pGKL2 (or K2), have become a model system to study such DNA elements. These VLEs have compact genomes with occasional overlaps of open reading frames (ORFs) and a high AT content of ~74% [5, 6]. The presence of both pGKL plasmids in several *K. lactis* strains is associated with the extensively studied yeast killer phenotype [7].

Functions of protein products for most ORFs encoded by the pGKL plasmids were predicted using bioinformatics approaches and some of these proteins were characterized by biochemical and genetic analyses [1]. Both pGKL1 and pGKL2 encode their own DNA polymerase and a terminal protein, and it is assumed that the mechanism of their replication is similar to the replication of viruses of the *Adenoviridae* family or *Bacillus subtilis* bacteriophage  $\phi$ 29 [8]. Consequently, marked sequence similarities between viral enzymes and putative products of several linear plasmid ORFs with expected function in replication and transcription resulted in yeast linear DNA plasmids being called virus-like elements nowadays [9]. Hence, it is believed that these VLEs may have originated from endosymbiotic bacteria or a virus [10]. Nevertheless, the exact evolutionary origin of the yeast linear VLEs remains unclear.

Transcription of VLE-specific genes has been shown to be independent of mitochondrial [11] and nuclear RNA polymerases [8, 12–14], and probably utilizes a VLE-specific RNA polymerase (RNAP). Experiments with bacterial reporter and yeast nuclear genes fused with pGKL-derived sequences identified an upstream conserved sequence (5'-ATNTGA-3') preceding each of the open reading frames. This upstream conserved sequence (UCS), usually located at a distance of 20–40 nucleotides prior to the start codon, is essential for cytoplasmatic transcription of the downstream located gene [15–17]. Sequences located farther upstream of the UCS element have been shown to have no effect on transcription [16]. The UCS element is highly conserved among all yeast VLEs and the UCS sequence derived from the *Pichia etchellsii* pPE1B element acts as a functional promoter when transplanted into the pGKL1 element [18]. Thus, the UCS element is a universal *cis*-acting component of the VLE-specific transcription system and it is essential for transcription initiation. After elongation, transcription then terminates after each gene because only monocistronic transcripts were revealed with Northern blot analyses of transcripts derived from ten pGKL-encoded ORFs [6, 12, 19–21]. This suggests the existence of a defined, yet unknown mechanism of transcription termination.

Unique RNAP subunits, and possibly also a putative helicase and the mRNA capping enzyme are the key elements of the VLE cytoplasmic transcription machinery. Protein products of *ORF6* (K2ORF6p; large subunit) and *ORF7* (K2ORF7p; small subunit) of the pGKL2 element should form a non-canonical RNAP. K2ORF6p was found to have a sequence similarity to three conserved regions of the two largest subunits ( $\beta$  and  $\beta'$  in bacteria) of canonical multisubunit RNAPs [22]. Sequence similarity of K2ORF6p to the  $\beta$  and  $\beta'$  subunit has recently been extended to 12 conserved regions shared by all bacterial, archaeal and eukaryotic

RNAPs [23], and the predicted structure of this enzyme thus resembles a fusion of the  $\beta$  subunit with a portion of the  $\beta'$  subunit. K2ORF7p was found to have sequence similarities to two conserved regions of the  $\beta'$  subunit, which are usually located at the C-terminus of  $\beta'$  [24].

The *ORF4* sequence of the pGKL2 element shows striking sequence similarity to viral helicases from the superfamily II of DEAD/H family helicases involved in transcription. The *K2ORF4* protein product (K2ORF4p) displays similarity with two Vaccinia virus helicases—(i) NPH-I, which is encoded by the *D11L* gene, and (ii) the small subunit of the heterodimeric Vaccinia virus early transcription factor (VETF) encoded by the *D6R* gene [22, 25]. NPH-I is known to provide the energy for elongation of transcription and for the release of RNA during transcription termination [26]. VETF functions as a transcription initiation factor that binds and bends the promoter region of early genes [27].

The protein product of *ORF3* (K2ORF3p) encoded by the pGKL2 element shows sequence similarity to the Vaccinia virus mRNA capping enzyme encoded by the *D1R* gene that consists of three domains responsible for the three enzymatic activities necessary to form the 5' mRNA cap structure [28]. The methyltransferase activity of the D1 protein of the poxvirus Vaccinia is allosterically stimulated by heterodimerization with a smaller protein encoded by the *D12L* gene [29, 30]. The complex of D1 and D12 proteins is sometimes also referred to as the vaccinia termination factor (VTF) because, together with NPH-I, it also acts as a transcription termination factor of early genes [31]. Triphosphatase and guanylyltransferase activities of K2ORF3p have been already confirmed experimentally *in vitro* [32].

As reported previously, the *K2ORF3*, *K2ORF4*, *K2ORF6* and *K2ORF7* genes are indispensable for the maintenance of the pGKL elements in the cell [19, 20, 24, 32]. However, understanding of interactions of their protein products with each other and with VLE DNA in the cell is lacking, as well as understanding of DNA sequence elements required for VLE transcription initiation and termination.

Here, we present a systematic *in vivo* study focusing on the architecture of the transcription complex of the yeast VLEs. Moreover, we identify a new promoter DNA element which is associated with 5' mRNA polyadenylation of most pGKL-encoded genes and we uncover a link between RNA stem loop structures and 3' end formation of VLE-specific mRNAs *in vivo*. Further, we present an extensive phylogenetic analysis of amino acid sequences of VLE RNAP subunits. Finally, we provide a detailed sequence analysis of pGKL promoters. Collectively, these analyses strongly suggest that the VLE transcription machinery has origin close to poxviruses.

## Results

### VLE RNAP subunits, mRNA capping enzyme, and helicase associate *in vivo*

To start characterizing the transcription machinery of the yeast VLEs we first tested whether the VLE RNAP subunits (K2ORF6p, K2ORF7p), the mRNA capping enzyme (K2ORF3p), and the putative helicase (K2ORF4p) form a complex *in vivo*.

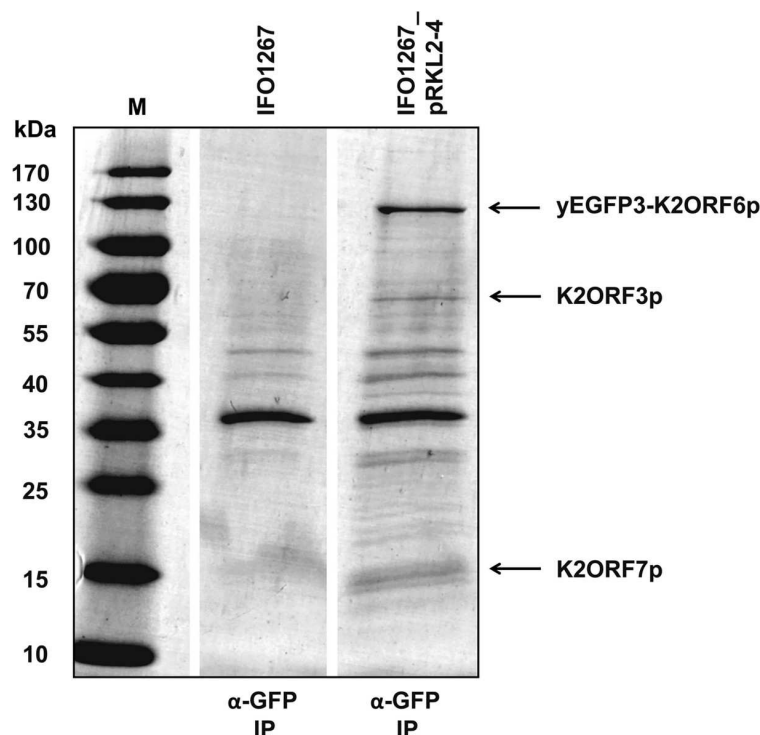
Initially, we tested interactions between K2ORF3p, K2ORF6p and K2ORF7p using a yeast two-hybrid system and its fluorescence variant called bimolecular fluorescence complementation but we failed to detect any interaction with either approach. This was most likely caused by the high AT content of VLE genes that was shown recently to impair their nuclear expression due to RNA fragmentation mediated by the polyadenylation machinery [33]. Therefore, we decided to prepare modified pGKL elements expressing the putative transcription machinery components containing various tags.

We prepared a strain encoding yeast enhanced green fluorescent protein 3 (yEGFP3) [34] that was fused to the N-terminus of the large RNAP subunit K2ORF6p (yEGFP3-K2ORF6p;

strain IFO1267\_pRKL2-4). We immunoprecipitated (IP) yEGFP3-K2ORF6p from this strain using GFP-Trap\_A agarose beads that contain a monoclonal antibody against common GFP variants. As a control, we used the wt strain IFO1267 that had no modifications. Extensive washing was used to remove weakly bound proteins. The bound proteins were eluted, resolved on SDS-PAGE, and stained with Coomassie Brilliant Blue G-250 (Fig 1). Gel lanes or bands of interest were excised from the gel, and analysed by mass spectrometry (MS). As shown in Table 1, peptides corresponding to yEGFP3-K2ORF6p, K2ORF7p (RNAP small subunit) and K2ORF3p (mRNA capping enzyme) were detected whereas no such peptides were identified in the parallel-treated IFO1267 control sample.

Then, to verify the interaction between the large RNAP subunit and the mRNA capping enzyme we decided to perform immunoprecipitation using tagged K2ORF3p as the bait. Hence, we prepared a strain encoding yEGFP3 fused to the C-terminus of K2ORF3p (IFO1267\_pRKL2-11 strain). Interestingly, selective cultivation of clones after transformation led to a loss of the pGKL1 element (S1 Fig). Immunoprecipitation using GFP-Trap\_A agarose beads and subsequent MS analysis revealed peptides corresponding to K2ORF6p, K2ORF7p, K2ORF4p, and K2ORF3p-yEGFP3 (S2 Fig).

To further validate the results, we prepared several strains containing combinations of these proteins with various tags: (i) strain IFO1267\_pRKL2-5 where yEGFP3-K2ORF6p was co-expressed together with K2ORF7p-FLAG, and control strain IFO1267\_pRKL2-15 that expressed only one tagged protein—K2ORF7p-FLAG; (ii) strain IFO1267\_pRKL2-6 co-expressing yEGFP3-K2ORF6p and K2ORF3p-HA, and control strain IFO1267\_pRKL2-14



**Fig 1. Identification of proteins associated with the large subunit of the VLE RNAP (K2ORF6p).** The gel shows Coomassie stained proteins affinity-purified with GFP-Trap\_A from strains IFO1267 (control) and IFO1267\_pRKL2-4 (containing yEGFP3-K2ORF6p). Proteins identified by mass spectrometry are indicated with arrows on the right side, and also listed in Table 1. M, protein molecular mass marker (PageRuler Prestained Protein Ladder, Fermentas); the respective molecular mass values are indicated on the left side.

<https://doi.org/10.1371/journal.ppat.1007377.g001>

Table 1. K2ORF6p-associated proteins.

Protein identified	MW (kDa)	Coverage	No. of peptides		Approach
			IFO1267_pRKL2-4	IFO1267	
yEGFP3-K2ORF6p	140.7	48%	170	0	whole lane analysis
K2ORF7p	15.5	53%	22	0	whole lane analysis
K2ORF3p	70.5	24%	13	-	band excision

IFO1267 (control) and IFO1267\_pRKL2-4 (yEGFP3-K2ORF6p) cells were grown to late exponential phase. The cells were lysed and yEGFP3-K2ORF6p was affinity-purified using GFP-Trap agarose beads. Bound proteins were eluted and resolved by SDS-PAGE, stained with Coomassie Brilliant Blue G-250, and analysed by mass spectrometry. The proteins identified, their molecular weight (MW), unique coverage, and the number of the detected peptides from both strains is listed.

<https://doi.org/10.1371/journal.ppat.1007377.t001>

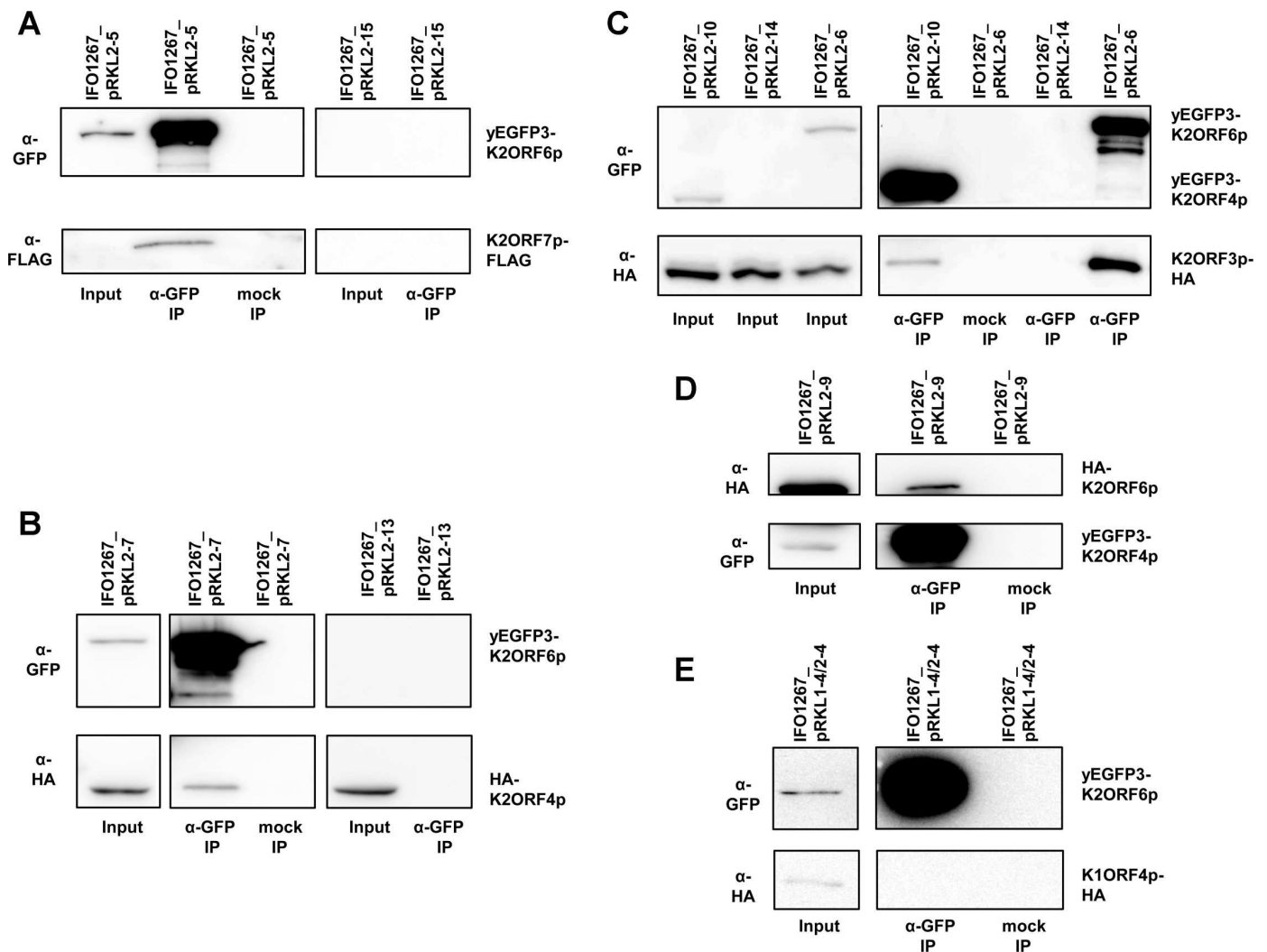
expressing only K2ORF3p-HA. The results of IP experiments followed with Western blotting clearly showed specific interactions between these proteins (Fig 2A and 2C).

Thus far, we had detected K2ORF3p, K2ORF6p, and K2ORF7p to be associated *in vivo*. For the fourth protein, the putative helicase K2ORF4p, the MS results suggesting it as part of the complex were not convincing due to low protein coverage (S2 Fig). To address whether it does, although perhaps weakly, interact with these proteins, we prepared a strain expressing HA-K2ORF4p together with yEGFP3-K2ORF6p (IFO1267\_pRKL2-7), and a control strain expressing HA-K2ORF4p only (IFO1267\_pRKL2-13). After IP and Western blotting we found HA-K2ORF4p to associate with yEGFP3-K2ORF6p (Fig 2B). The results clearly showed that the putative helicase was specifically associated with the large RNAP subunit (or, was present in a complex containing this subunit) because yEGFP3-K2ORF6p and HA-K2ORF4p were not bound to the empty agarose beads, and HA-K2ORF4p alone did not bind to the GFP-Trap antibody (Fig 2B).

Because the association of the putative helicase with the large RNAP subunit seemed rather weak, we decided to perform reciprocal immunoprecipitation. Further, we also tested whether the mRNA capping enzyme associated with the putative helicase. Strains expressing (i) yEGFP3-K2ORF4p together with HA-K2ORF6p (IFO1267\_pRKL2-9), and (ii) yEGFP3-K2ORF4p together with K2ORF3p-HA (IFO1267\_pRKL2-10) were prepared. With the first combination we confirmed that HA-K2ORF6p associated with yEGFP3-K2ORF4p (Fig 2D). With the second combination we found that K2ORF3p-HA associated with yEGFP3-K2ORF4p (Fig 2C). The detected interactions were specific because yEGFP3-K2ORF4p and HA-K2ORF6p were not bound to the empty agarose beads and K2ORF3p-HA alone did not bind to the GFP-Trap antibody (Fig 2C and 2D).

As an additional control to demonstrate that the observed interactions were specific, we used another pGKL-encoded protein with a function unrelated to transcription. We selected K1ORF4p, a subunit of the toxin responsible for VLE-associated killer yeast phenotype [35]. A strain co-expressing yEGFP3-K2ORF6p together with K1ORF4p-HA (IFO1267\_pRKL1-4/2-4) was prepared. We found that K1ORF4p-HA was not associated with yEGFP3-K2ORF6p (Fig 2E).

Finally, we wanted to know whether the association of the putative helicase and mRNA capping enzyme with the large RNAP subunit was dependent on nucleic acids. We prepared lysates from IFO1267\_pRKL2-6 (yEGFP3-K2ORF6p, K2ORF3p-HA) and IFO1267\_pRKL2-7 (yEGFP3-K2ORF6p, HA-K2ORF4p) strains. The lysates were incubated with GFP-Trap\_A beads and, after washing, the beads were split into two parts which were treated or mock-treated with Benzonase Nuclease to digest DNA and RNA. Then, the beads were again extensively washed and the bound proteins were eluted. Subsequent Western blot analysis revealed both K2ORF3p-HA and HA-K2ORF4p to associate with yEGFP3-K2ORF6p regardless of the presence or absence of nucleic acids which was confirmed by PCR (S3 Fig).



**Fig 2. Interactions between RNAP subunits, mRNA capping enzyme, and putative helicase of the yeast VLEs.** (A) Western blot of immunoprecipitated ( $\alpha$ -GFP IP) and mock immunoprecipitated (mock IP) proteins from IFO1267\_pRKL2-5 (yEGFP3-K2ORF6p, K2ORF7p-FLAG) or control IFO1267\_pRKL2-15 cells (K2ORF7p-FLAG), respectively. The strains used are indicated above the lanes. The antibodies used for Western blot are indicated on the left hand side of the strips. The proteins detected are indicated on the right hand side of the strips. Positions of the identified proteins corresponded with theoretical molecular weight of the full length recombinant proteins, as determined by positions of the protein mass markers. Input represented approximately 1/100 of the sample and IP represented approximately 1/2 of the sample in this and the other immunoprecipitation experiments. Mock immunoprecipitations in all experiments were done using empty agarose beads. The same experimental scheme is used throughout this figure. (B) Western blot analysis of immunoprecipitations from lysates from IFO1267\_pRKL2-7 (yEGFP3-K2ORF6p, HA-K2ORF4p) and control IFO1267\_pRKL2-13 (HA-K2ORF4p) cells (indicated above the lanes). The ( $\alpha$ -GFP) and anti-HA ( $\alpha$ -HA) antibodies used are indicated on the left hand side; the detected proteins on the right hand side. (C) Western blot analysis of immunoprecipitations from lysates from strains IFO1267\_pRKL2-6 (yEGFP3-K2ORF6p, K2ORF3p-HA), IFO1267\_pRKL2-10 (yEGFP3-K2ORF4p, K2ORF3p-HA), and IFO1267\_pRKL2-14 (control). (D) Western blot analysis of immunoprecipitations from lysates from IFO1267\_pRKL2-9 cells (yEGFP3-K2ORF4p, HA-K2ORF6p). (E) Western blot analysis of immunoprecipitations from lysates from IFO1267\_pRKL1-4/2-4 cells (yEGFP3-K2ORF6p, K1ORF4p-HA).

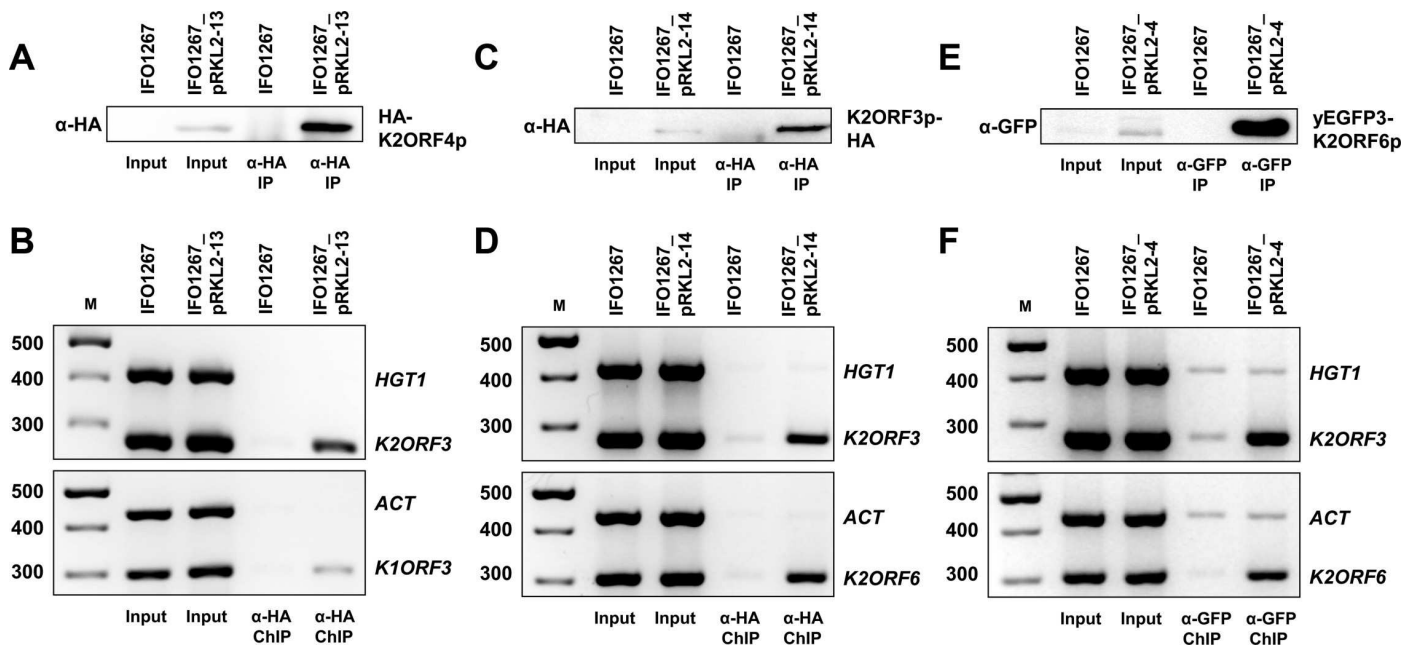
<https://doi.org/10.1371/journal.ppat.1007377.g002>

Taken together, the immunoprecipitation, mass spectrometry, and Western blot results demonstrated the existence of the hypothesized VLE-specific transcription complex because K2ORF3p, K2ORF4p, K2ORF6p, and K2ORF7p were specifically associated *in vivo*. This association was independent of nucleic acids. Finally, K2ORF3p, K2ORF6p, and K2ORF7p appeared to form a core transcription complex with relatively strong mutual interactions to which K2ORF4p bound relatively weakly.



## RNAP, mRNA capping enzyme, and helicase are associated with VLE-specific DNA *in vivo*

Previous results confirmed the existence of the hypothesized VLE transcription complex *in vivo*. It is believed that this transcription complex is VLE-specific [22]. Therefore, it should possibly interact only with the VLE DNA *in vivo*, which was never formally tested. Hence, we performed *in vivo* chromatin immunoprecipitation. We used the IFO1267\_pRKL2-13 strain expressing HA-K2ORF4p, IFO1267\_pRKL2-14 strain expressing K2ORF3p-HA, IFO1267\_pRKL2-4 strain expressing yEGFP3-K2ORF6p, and the IFO1267 control strain. First, we verified that the mouse monoclonal anti-HA HA-7 agarose efficiently immunoprecipitated HA-K2ORF4p and K2ORF3p-HA (Fig 3A and 3C), and that the GFP-Trap\_A agarose beads efficiently immunoprecipitated yEGFP3-K2ORF6p (Fig 3E). Then, we performed chromatin immunoprecipitation of HA-K2ORF4p, K2ORF3p-HA, and yEGFP3-K2ORF6p from formaldehyde cross-linked cells. The immunoprecipitated and input DNAs were used as templates for subsequent PCR analysis using primers designed to detect chromosomal or VLE DNA. We used primers specific for *K. lactis* actin (*ACT*) and high-affinity glucose transporter (*HGT1*) genes as markers of chromosomal DNA, and toxin immunity (*K1ORF3*), mRNA capping enzyme (*K2ORF3*), and large RNAP subunit (*K2ORF6*) genes as markers of pGKL elements. We found, that HA-K2ORF4p, K2ORF3p-HA, and yEGFP3-K2ORF6p were specifically associated with pGKL elements and not with



**Fig 3. Physical association of the putative helicase (K2ORF4p), mRNA capping enzyme (K2ORF3p), and the large RNAP subunit (K2ORF6p) of the yeast VLEs with VLE-specific DNA.** (A) Western blot of HA-K2ORF4p that was affinity-purified from lysates of IFO1267\_pRKL2-13 (HA-K2ORF4p) and IFO1267 (control) cells. The strains used are indicated above the lanes. The antibody used is indicated on the left hand side of the strip. The protein detected is indicated on the right hand side of the strip. (B) PCR analysis of the presence of chromosomal (*ACT*, *HGT1*) or VLE (*K1ORF3*, *K2ORF3*) DNA in chromatin immunoprecipitated using anti-HA HA-7 agarose from IFO1267\_pRKL2-13 (HA-K2ORF4p) and IFO1267 (control) cells. Samples of individually performed gene-specific PCRs were analysed in 2.5% agarose gel stained with ethidium bromide. The identity of the bands (genes) is indicated on the right. M, DNA molecular mass marker (GeneRuler 100 bp Plus DNA Ladder, Fermentas). The respective values are indicated on the left. (C) Western blot of K2ORF3p-HA that was affinity-purified from lysates of IFO1267\_pRKL2-14 (K2ORF3p-HA) and IFO1267 (control) cells. (D) PCR analysis of the presence of chromosomal (*ACT*, *HGT1*) or VLE (*K2ORF3*, *K2ORF6*) DNA in chromatin immunoprecipitated using anti-HA HA-7 agarose from IFO1267\_pRKL2-14 and IFO1267 cells. (E) Western blot of yEGFP3-K2ORF6p that was affinity-purified from lysates of IFO1267\_pRKL2-4 (yEGFP3-K2ORF6p) and IFO1267 (control) cells. (F) PCR analysis of the presence of chromosomal (*ACT*, *HGT1*) or VLE (*K2ORF3*, *K2ORF6*) DNA in chromatin immunoprecipitated using GFP-Trap agarose beads from IFO1267\_pRKL2-4 and IFO1267 cells.

<https://doi.org/10.1371/journal.ppat.1007377.g003>



chromosomal DNA (Fig 3B, 3D and 3F). These results were also confirmed by semiquantitative real-time PCR.

We concluded that the mRNA capping enzyme, and previously uncharacterized RNAP and helicase were associated with VLE-specific DNA *in vivo*, which further supports involvement of these proteins in transcription of VLEs.

### Slippage of RNAP at the initiation site results in 5' polyadenylation of VLE mRNAs

Next, we wished to characterize transcription initiation of the VLE genes. Our previous 5' RACE-PCR experiments had revealed 5' cap structures on the VLE-specific mRNAs, likely synthesized by VLE-encoded K2ORF3p mRNA capping enzyme, and also the presence of non-templated 5' poly(A) leaders of heterogeneous lengths in mRNAs of 12 pGKL genes except for *K2ORF2*, *K2ORF3* and *K2ORF8* [36]. Interestingly, heterogeneous 5' poly(A) leaders are a known feature of poxviral intermediate and late transcripts [37, 38]. It was shown that the 5' poly(A) leader was produced by slippage of Vaccinia virus RNAP on three consecutive thymidine residues in the template strand at the initiator region (INR) where transcription starts both *in vivo* and *in vitro* [39, 40].

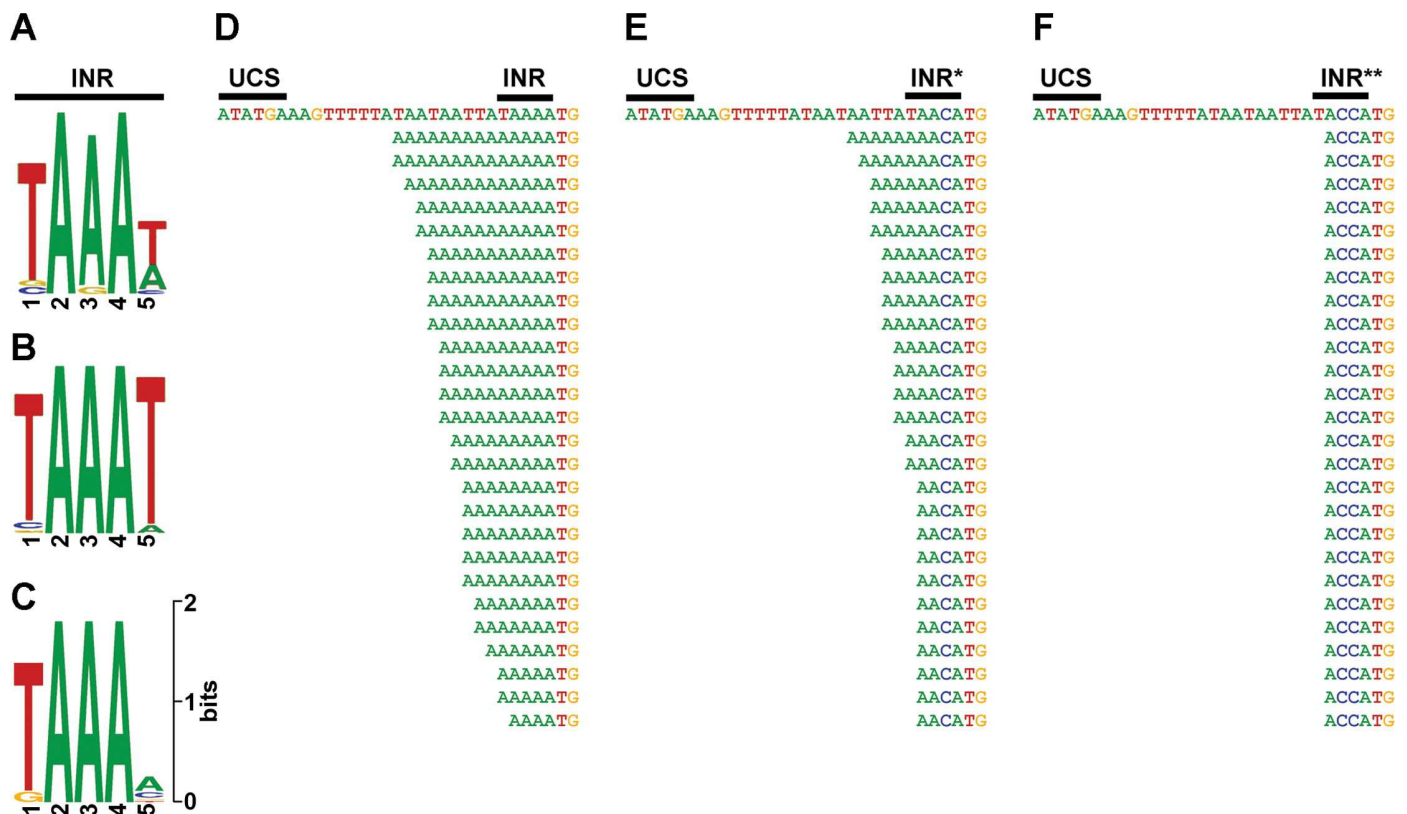
Fig 4A and 4B show the sequence logo of the INR consensus motif (TAAAT) of Vaccinia virus intermediate and late genes, respectively [41]. Interestingly, we were able to locate a similar INR-like consensus sequence (TAAAN) in promoters of all 12 pGKL genes whose transcripts were 5' polyadenylated (Fig 4C). For the transcription start site (TSS) annotation, the first adenosine residue of the motif was considered to encode the initiating nucleoside triphosphate.

Subsequently, we tested whether the putative INR was responsible for the 5' end polyadenylation of the pGKL-derived transcripts. We prepared three *K. lactis* strains with modified pGKL1 elements encoding the G418 resistance marker under the control of the K1UCR2 promoter. We prepared three variants of the K1UCR2 promoter that differed in the INR sequence: (i) TAAAA (wt; strain IFO1267\_pRKL1-1); (ii) TAACA (strain IFO1267\_pRKL1-2); and (iii) TACCA (strain IFO1267\_pRKL1-3). Then, we purified total RNA from the three strains, prepared cDNA, and performed 5' RACE-PCR to determine the 5' end sequences. The results showed that the 5' poly(A) leader was present when the K1UCR2 sequence contained the putative wt INR (TAAAA INR) (Fig 4D). When TAACA INR\* was used, the length of the 5' end poly(A) was significantly reduced (Fig 4E). When TACCA INR\*\* was used, the poly(A) leader disappeared altogether (Fig 4F).

Therefore, we concluded that slippage of VLE RNAP at the initiation site was the mechanism responsible for 5' polyadenylation of the transcripts. Moreover, the identified INR sequence constituted an independent DNA element, not influenced by the sequence of the gene because the pattern of the sequenced 5' RACE-PCR clones for *K1ORF2* transcripts was the same as for *G418<sup>R</sup>* transcripts produced from the K1UCR2 with the wt INR [36].

### RNA stem loop structures influence 3' end formation of VLE-specific mRNAs *in vivo*

Our previous 3' RACE-PCR experiments had revealed the absence of 3' poly(A) tails in mRNAs of all 15 pGKL ORFs [36]. To shed light on the mode of transcription termination of the VLEs, we tried to identify sequence/secondary structure elements/signals near the 3' termini. First, we searched for sequence motifs within the last 150 nt of each transcript that would be shared among the 15 pGKL ORFs but we detected none. Second, we searched for secondary structure motifs using the RNAstructure Server [42]. We identified putative RNA

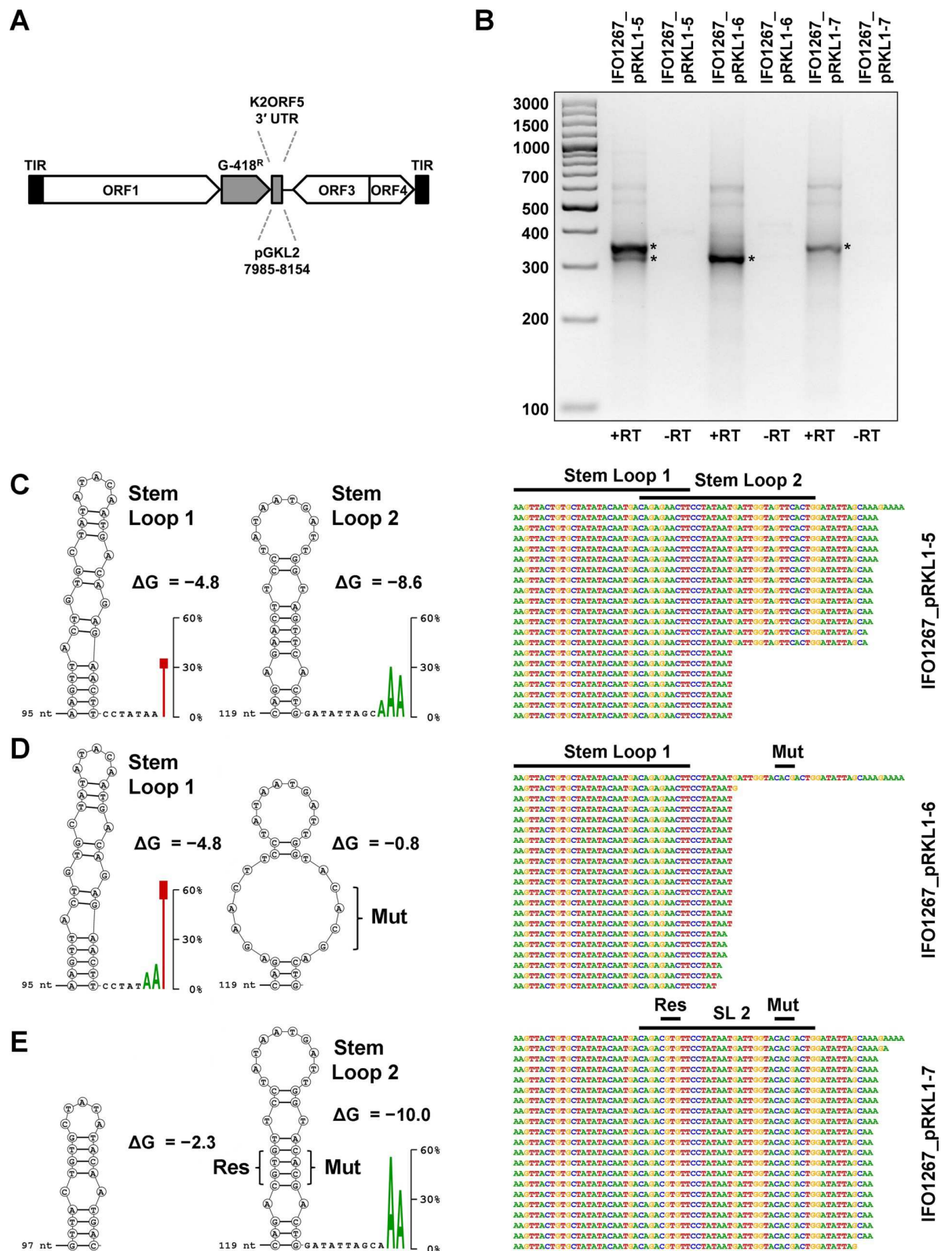


**Fig 4. Promoters of yeast VLEs contain initiator region (INR) responsible for non-templated 5' polyadenylation of mRNAs.** (A) Sequence logo of the initiator region in promoter sequences of Vaccinia virus intermediate genes [41]. (B) Sequence logo of the initiator region in promoter sequences of Vaccinia virus late genes [41]. (C) Sequence logo of the putative INR identified in promoters of 12 ORFs encoded by pGKL elements. (D) 5' RACE-PCR analysis of the *G418<sup>R</sup>* gene from the IFO1267\_pRKL1-1 strain. In this and the following panels, the upper sequence corresponds to the template (plasmid) DNA and the UCS is indicated; sequences situated below represent individual sequenced cDNA clones (the 5' untranslated region is shown in full till the translation start codon, ATG). Guanosine residues corresponding to the original 5' mRNA caps which were present in some of the cDNA clones are omitted in this representation for clarity. (E) 5' RACE-PCR analysis of the *G418<sup>R</sup>* gene from the IFO1267\_pRKL1-2 strain bearing a promoter mutation in the putative INR reducing the number of consecutive adenosine residues in the template. (F) 5' RACE-PCR analysis of the *G418<sup>R</sup>* gene from the IFO1267\_pRKL1-3 strain bearing promoter mutations in the putative INR abolishing consecutive adenosine residues in the template.

<https://doi.org/10.1371/journal.ppat.1007377.g004>

stem loop structures close to the experimentally determined 3' ends of cDNA (S4 Fig). The putative RNA stem loops were typically in the vicinity of the respective ORF's stop codon with the median distance of 26 nt, and Gibbs free energy of  $-7.5$  kcal/mol.

Hence, we tested, whether the predicted RNA stem loop structures influenced the 3' mRNA end formation. Because pGKL elements contain almost no intergenic regions, the putative RNA stem loops are localized in the coding sequences of adjacent ORFs or within the terminal inverted repeats. This means that their sequences cannot be subjected to mutagenesis without the possibility of altering VLE functions. Therefore, we prepared a *K. lactis* strain with a modified pGKL1 element encoding the *G418* resistance marker under control of K1UCR2, followed by the 3' UTR of the *K2ORF5* gene (strain IFO1267\_pRKL1-5; Fig 5A). The distal part of the *K2ORF5* 3' UTR contained two putative partially overlapping RNA stem loops termed Stem loop 1 and 2 (S4I Fig). 3' RACE-PCR experiments revealed a transcription termination pattern that could be attributed to the presence of both Stem loop 1 and 2 (Fig 5B and 5C). Next, using the same promoter-gene-3' UTR arrangement, we prepared a strain with 4 nucleotide mutations destabilizing the base pairing in the middle of the putative Stem loop 2 (strain IFO1267\_pRKL1-



**Fig 5. RNA stem loop structures influence the 3' end formation of VLE-specific mRNAs *in vivo*.** (A) Schematic representation of recombinant pGKL1 elements where the *G418<sup>R</sup>* gene is followed by the coding sequence of wild-type (pRKL1-5) or modified (pRKL1-6, pRKL1-7) 3' untranslated region of the *K2ORF5* gene. TIR—terminal inverted repeat. (B) 3' RACE-PCR analysis of individual mRNAs corresponding to the *G418<sup>R</sup>* gene expressed from modified pGKL1 elements. Samples were analyzed in 3.0% agarose gel stained by ethidium bromide. The strains used to purify the RNA are indicated above the lanes. M, DNA molecular mass marker (GeneRuler 100 bp Plus DNA Ladder, Fermentas). The respective values are indicated on the left. Specific products that were cloned to the pCR4-TOPO vector and used for sequencing are labelled with asterisks. Reverse transcription was carried out in the presence (+RT) and absence (-RT) of reverse transcriptase. (C) 3' RACE-PCR analysis of the *G418<sup>R</sup>* gene from IFO1267\_pRKL1-5 strain. In this and the following panels, the upper sequences on the right correspond to the template (plasmid) DNA; sequences situated below represent 3' end regions of individual sequenced cDNA clones. Positions of the putative RNA stem loops are indicated above the sequences. Predicted RNA stem loops are displayed as cDNA nucleotide letters in circles on the left and the values of Gibbs free energy ( $\Delta G$ ) in kcal/mol are displayed for each structure. Stem loop distances from the gene stop codon are shown as numbers of nucleotides. The last few 3' end nucleotides of the experimentally determined 3' ends of cDNA are shown as letters enlarged proportionally to their occurrence (in %) in the sequenced clones in the case when these nucleotides were detected in at least two independent clones. (D) 3' RACE-PCR analysis of the *G418<sup>R</sup>* gene from the IFO1267\_pRKL1-6 strain. Mut, the mutated stem loop. (E) 3' RACE-PCR analysis of the *G418<sup>R</sup>* gene from IFO1267\_pRKL1-7 strain. Mut, the mutated stem loop; Res, the rescued stem loop.

<https://doi.org/10.1371/journal.ppat.1007377.g005>

6). For this construct, we detected a transcription termination pattern that could be attributed to the presence of only Stem loop 1 (Fig 5B and 5D). Subsequently, we prepared a strain (IFO1267\_pRKL1-7) where we changed the sequence but not the base-pairing of 4 nucleotides within the Stem loop 2. 3' RACE-PCR experiments revealed a transcription termination pattern that could be attributed to the presence of the rescued Stem loop 2 (Fig 5B and 5E). Note that the rescue mutations distinctly altered the length and the Gibbs free energy (destabilizing Stem loop 1) of the overlapping Stem loop 1 and this was likely the reason why transcription termination from the Stem loop 1 was not detected in the IFO1267\_pRKL1-7 strain (Fig 5B and 5E).

Finally, we determined whether promoter sequences or coding sequences of a gene could affect the 3' mRNA end formation. We used *K2ORF5*, *G418*, and hygromycin B resistance genes under the control of *K2UCR5*, *K1UCR2* and *K1UCR3*, respectively, located on the pGKL2 element. Downstream of the coding sequence of each of these genes we positioned the 3' untranslated region (UTR) of the *K2ORF5* gene. We purified total RNA from the respective *K. lactis* strains, prepared cDNA, and performed 3' RACE-PCR experiments. Although the coding sequences of the aforementioned genes differed both in length and AT content, the pattern of their 3' termini was highly similar (S4 and S5 Figs).

We concluded that RNA stem loop structures were essential for the 3' end formation of VLE-specific mRNAs *in vivo*, presumably acting as factor-independent intrinsic terminators. Moreover, this termination was independent of the promoter and the gene used both with respect to its sequence and length.

### VLE-specific RNA polymerase has unique architecture

To facilitate interpretation of the experimental data we created a 3D model of the pGKL-specific RNAP. This was feasible due to sequence similarity between parts of *K2ORF6p*, *K2ORF7p* and conserved regions of the canonical multisubunit RNAPs [20, 23]. The structural models of the two subunits covered 92.7% of the *K2ORF6p* sequence and 62.1% of the *K2ORF7p* sequence, respectively.

Superimposition of the models over the *S. cerevisiae* RNAP II elongation complex (PDB ID: 2NVQ) using DaliLite v. 3 showed that all known regions of *K2ORF6p* and *K2ORF7p* with sequence similarity to conserved regions of the canonical RNAPs were modelled accordingly. Interestingly, parts of *K2ORF6p* were modelled by  $\beta$ a1,  $\beta$ a6,  $\beta$ a13, and  $\beta$ a16 conserved regions, which were not detected previously to be present in *K2ORF6p*. We verified the proper model: template alignment of these regions by manually constructed sequence alignments with other canonical RNAPs and these alignments indeed showed sequence similarities between *K2ORF6p* and the aforementioned regions (S6 Fig).



The overall distribution of the conserved regions within K2ORF6p and K2ORF7p is depicted in Fig 6A. It should be noted that K2ORF6p displayed a unique fusion between  $\beta$  and  $\beta'$  subunit conserved regions, which is not known to be present in any other canonical or non-canonical RNAP. This fusion seemed to be essential for the maintenance of pGKL elements in yeast—a VLE where we divided *K2ORF6* into two genes, based on their homology to  $\beta$  and  $\beta'$ , was unable to substitute wt pGKL2 in the cell. On the other hand, a division of the  $\beta'$  subunit (between  $\beta'a15$  and  $\beta'a16$ ) into two polypeptides, similarly as in K2ORF6p and K2ORF7p, is present in RNAPs of some *Archaeal* species [43].

Fig 6B and 6C show the *S. cerevisiae* RNAP II Rpb2 subunit ( $\beta$  subunit homolog) and the K2ORF6p model, respectively. It is clear that almost all conserved  $\beta$  subunit regions are present in K2ORF6p, and only the spacing between them is shorter. Fig 6D and 6E show the *S. cerevisiae* RNAP II Rpb1 subunit ( $\beta'$  subunit homolog) and relevant homologous portions of K2ORF6p/K2ORF7p subunits, respectively. Remarkably, more than half of the  $\beta'$  subunit conserved regions, such as most of the clamp domain ( $\beta'a1 - \beta'a10$  regions) and secondary-channel rim helices ( $\beta'a13, \beta'a14$  regions), are missing in VLE RNAP (Fig 6D).

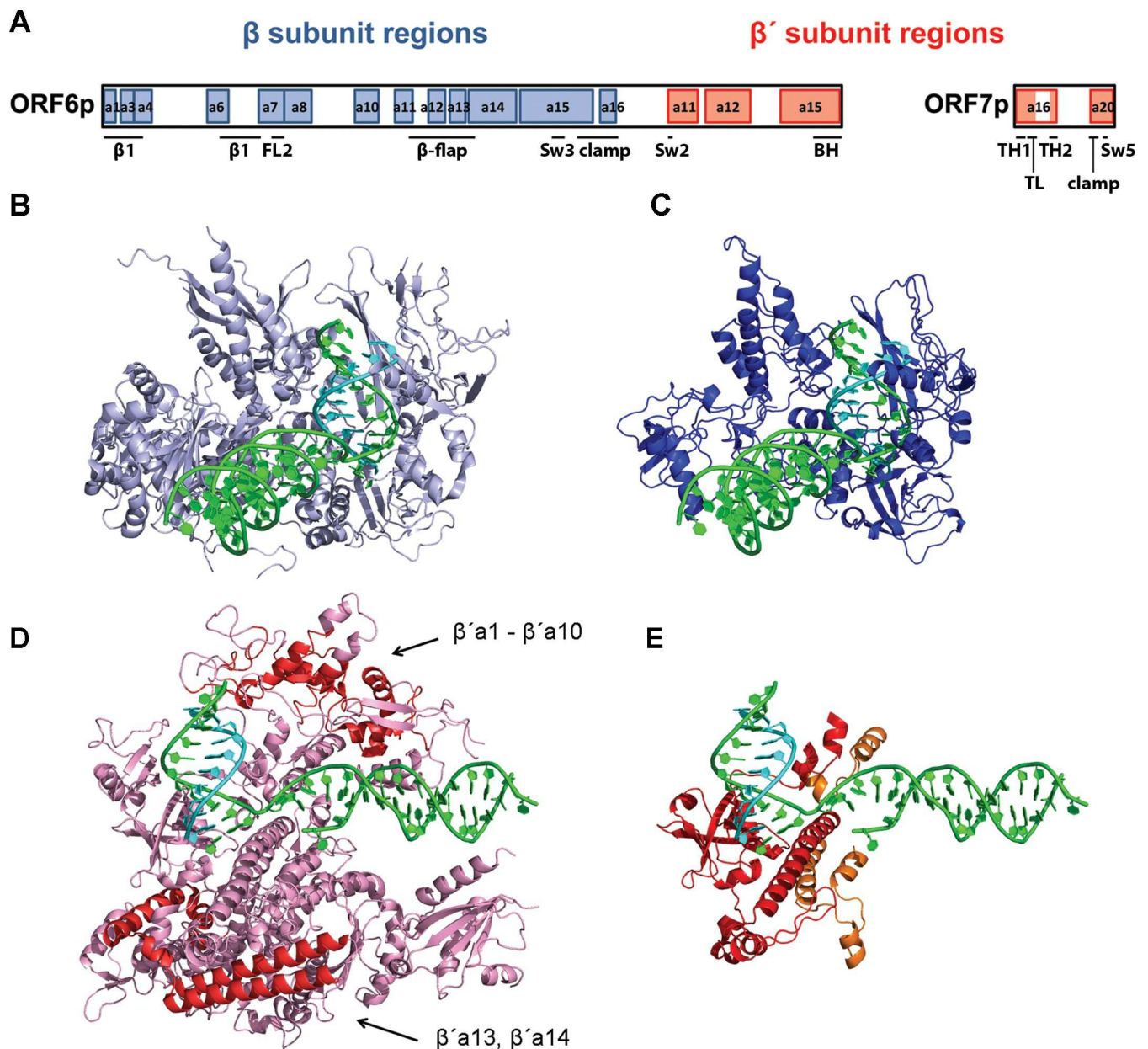
We concluded that the VLE RNAP displayed a unique and novel architecture.

### VLE-specific RNA polymerase has a viral origin

An extensive phylogenetic analysis of yeast VLE RNAPs has not been performed yet. Earlier analyses focused only on conserved regions  $\beta'a11$  and  $\beta'a12$  of K2ORF6p. These studies established that VLE RNAP belongs to multisubunit RNAPs rather than to single-subunit RNAPs encoded by mitochondrial linear plasmids of fungi and plants [10, 44]. Hence, it is believed that the *ORF6* and *ORF7* genes of VLEs were derived from eukaryotic multisubunit RNAP genes of ancestral host cells [1, 23] or, alternatively, that the *ORF6* gene is an ancient representative of multisubunit RNAP diversification from times when  $\beta$  and  $\beta'$  constituted a single protein [45].

We performed a detailed phylogenetic analysis to delve deep into the evolutionary past of yeast VLEs. We used sequences of  $\beta'a11$ - $\beta'a12$ ,  $\beta'a15$ - $\beta'a16$  and  $\beta'a20$  conserved regions of ORF6 and ORF7 proteins from all sequenced yeast VLEs. Alignment containing these conserved regions was then combined with the published alignment of  $\beta'$  subunit conserved regions of canonical multisubunit RNAPs [46] and was used to construct a maximum likelihood phylogenetic tree.

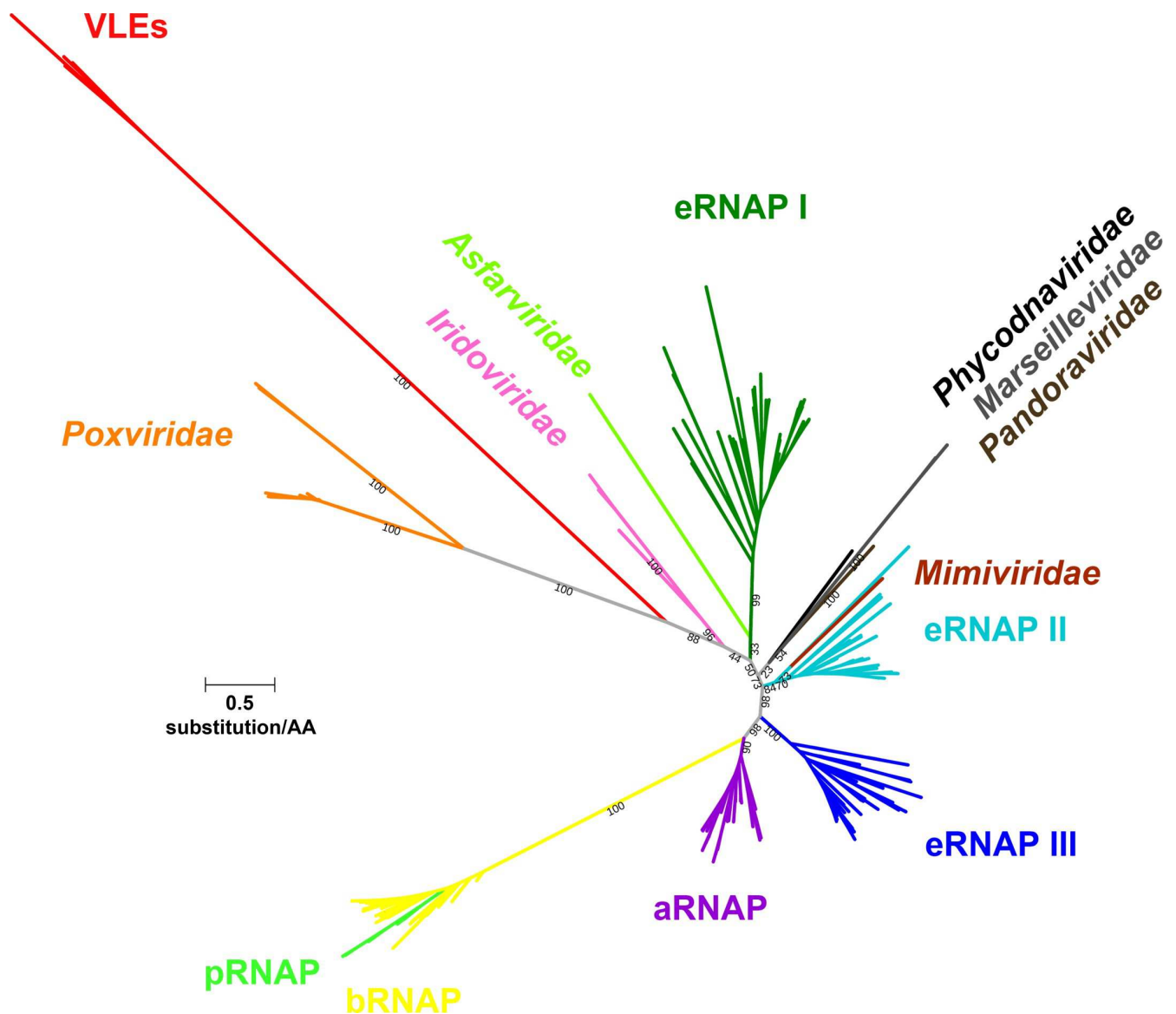
Fig 7 shows the unconstrained phylogenetic tree for conserved  $\beta'$  subunit regions. The tree surprisingly suggests monophyly of VLE RNAPs with viral RNAPs of the *Poxviridae* family, with viral RNAPs of the *Iridoviridae* family also belonging to the same clade. It is believed that ancestral  $\beta'$  subunit orthologs of all nucleo-cytoplasmic large DNA viruses (NCLDVs) were monophyletic, but the ancestral gene was displaced in Asfarvirus and Mimivirus for eukaryotic RNAP I and RNAP II gene, respectively [47]. In our analysis, Asfarvirus RNAP clustered with eukaryotic RNAPs I, and Mimivirus RNAP clustered with eukaryotic RNAPs II (Fig 7), which is consistent with that view. Interestingly, a single NCLDV clade (with the exception of *Asfarviridae* and *Mimiviridae*) was not recovered in our unconstrained tree. Further, we tested alternative evolutionary hypotheses for VLE RNAPs via constrained tree topology approach where we enforced monophyly of VLE RNAPs with other viral and cellular RNAPs, and we compared the likelihood of the original tree with the likelihood of the constrained trees. From these analyses, it was apparent that the likelihood of the unconstrained tree suggesting monophyly of the yeast VLE RNAP with poxviral RNAP was the best, although monophyly of VLE RNAP with RNAP of other NCLDVs could not be rejected by statistical tests (S7 Fig). Importantly, monophyly of VLE RNAP with eukaryotic RNAP I and RNAP II was rejected at a statistically significant level based on Expected Likelihood Weight test (S7 Fig).



**Fig 6. Predicted 3D structure of yeast VLE non-canonical RNAP encoded by the *ORF6* and *ORF7* genes.** (A) Schematic representation of the primary sequence of pGKL-encoded RNAP showing similarity to conserved regions of the catalytic subunits of canonical multisubunit RNAPs.  $\beta$  subunit conserved regions (blue) and  $\beta'$  subunit conserved regions (red) present in K2ORF6p and K2ORF7p are drawn to scale. Conserved regions are named according to ref. [46]. Sequence alignments for newly detected similarity of *ORF6* protein products to  $\beta$ a1,  $\beta$ a6,  $\beta$ a13 and  $\beta$ a16 conserved regions are provided in S6 Fig. BH, Bridge helix; FL2, Fork loop 2; Sw2, Switch 2; Sw3, Switch 3; Sw5, Switch 5; TH1, Trigger helix 1; TH2, Trigger helix 2; TL, Trigger loop. (B) 3D crystal structure of *Saccharomyces cerevisiae* RNA polymerase II elongation complex showing the Rpb2 subunit ( $\beta$  subunit homolog, light blue), DNA (green) and RNA (cyan). This figure is based on 2NVQ. (C) 3D model of pGKL RNAP showing K2ORF6p residues 1–693 ( $\beta$  subunit homolog, blue), DNA (green) and RNA (cyan). Nucleic acids in this as well as the following structures are based on 2NVQ. The structural similarity between the pGKL RNAP  $\beta$  subunit and RNAP II Rpb2 as calculated by DaliLite v. 3 is a root-mean-square deviation (RMSD) of 3.9 Å over 655 aligned C $\alpha$  positions, 23% sequence identity, and the TM-score 0.87. (D) 3D crystal structure of *Saccharomyces cerevisiae* RNAP II elongation complex showing the Rpb1 subunit ( $\beta'$  subunit homolog, pink), DNA (green) and RNA (cyan). Arrows indicate  $\beta'$  regions shared by multisubunit RNAPs (red) that are clearly missing in the VLE RNAP. This figure is based on 2NVQ. (E) 3D model of pGKL RNAP showing K2ORF6p residues 754–882 and 894–974 ( $\beta'$  subunit homolog, red), K2ORF7p residues 1–52 and 103–132 ( $\beta'$  subunit homolog, orange), DNA (green) and RNA (cyan). The structural similarity between pGKL RNAP  $\beta'$  subunit and RNAP II Rpb1 as calculated by DaliLite v. 3 is a RMSD of 2.3 Å over 123 aligned C $\alpha$  positions, 31% sequence identity, and the TM-score 0.89 for K2ORF6p residues 754–882; RMSD of 1.4 Å over 80 aligned C $\alpha$  positions, 21% sequence identity, and the TM-score 0.86 for K2ORF6p residues 894–974; RMSD of 2.9 Å over 43 aligned C $\alpha$  positions, 23% sequence identity, and the TM-score 0.57 for K2ORF7p residues 1–52; RMSD of 1.3 Å over 29 aligned C $\alpha$  positions, 13% sequence identity, and the TM-score 0.56 for K2ORF7p residues 103–132. For details concerning structure modelling see [Materials and methods](#).

<https://doi.org/10.1371/journal.ppat.1007377.g006>





**Fig 7. Phylogenetic analysis of yeast VLE RNA polymerases with canonical RNAPs.** A phylogram of  $\beta'$  subunit homologs based on amino acid sequence alignment of  $\beta'$  subunit conserved regions of selected canonical RNAPs and those  $\beta'$  subunit conserved regions present in *ORF6* and *ORF7* genes of the yeast VLEs. The maximum likelihood unconstrained tree is displayed as an unrooted phylogram where the length of the branches is proportional to the calculated evolutionary distance of individual sequences. Leaves defining the different classes of RNAPs are labeled. Used abbreviations for RNAP groups: aRNAP, archaeal RNAP; bRNAP, bacterial RNAP; eRNAP, eukaryotic RNAP; pRNAP, plastid RNAP. Length scale of branches of an average value of 0.5 substitution per amino acid residue is shown as a line near the tree. Selected branch support values calculated from 1 000 ultrafast bootstrap replicates optimized using nearest neighbour interchange (NNI) to reduce overestimating support are indicated in black. For details concerning phylogenetic analysis see [Materials and methods](#).

<https://doi.org/10.1371/journal.ppat.1007377.g007>

Taken together, our phylogenetic analysis surprisingly points to a viral origin of VLE RNAPs, possibly close to poxviruses, rather than to a cellular origin, contradicting previous hypotheses about the origin of these enzymes [1, 23, 45].

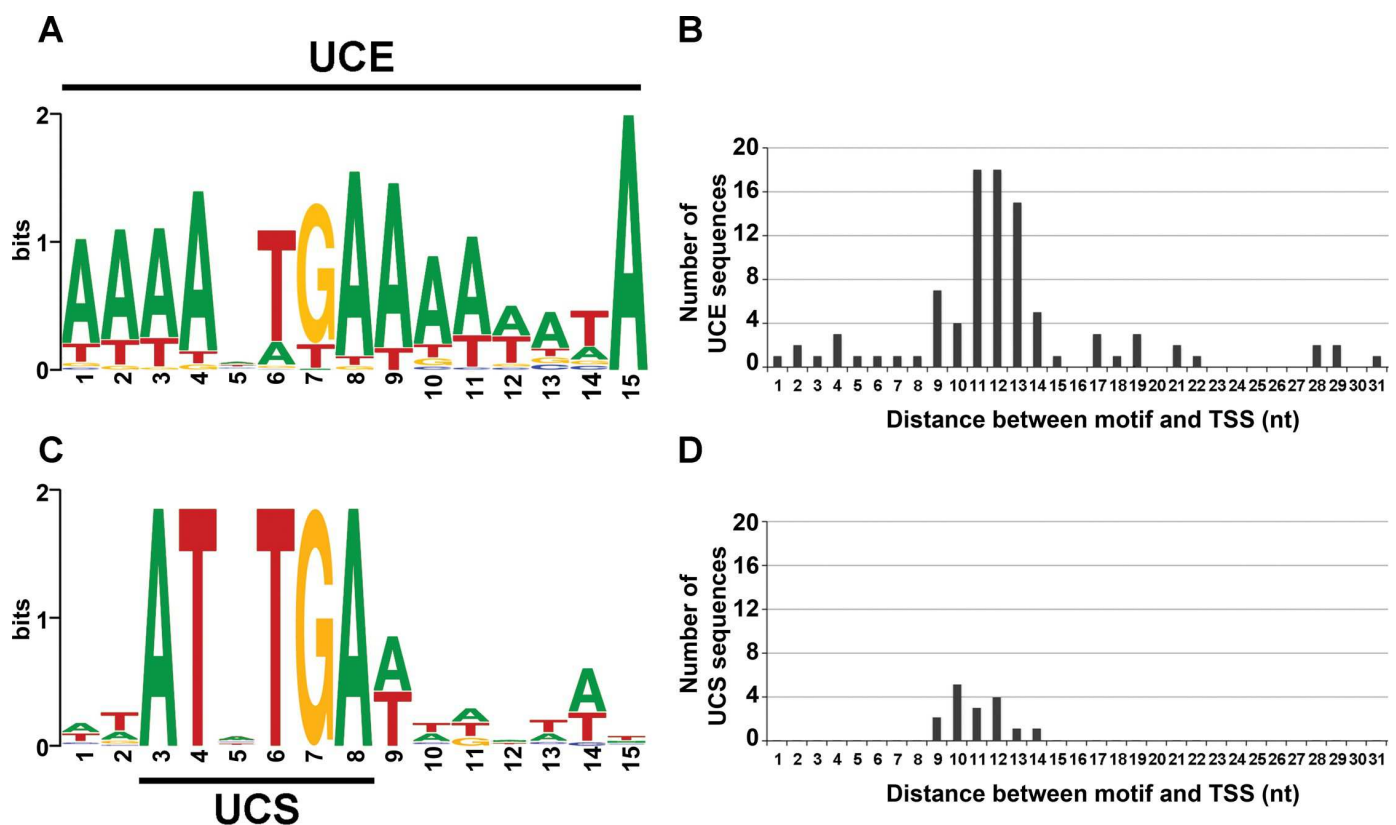
### VLE promoters have a viral origin

Although the UCS (5'-ATNTGA-3') essential for transcription of VLEs was identified a long time ago, no similarities with known promoters that would indicate its origin were reported. We

extended the UCSs preceding all pGKL-encoded ORFs both upstream and downstream by ~10 bp, and we created a consensus motif. This consensus was then used to search for similar elements that were associated with transcription by multisubunit RNAPs. We particularly focused on promoters of viral RNAPs because our phylogenetic analysis of VLE RNAPs had suggested a viral origin.

Notably, we detected great sequence similarity between the extended UCS (Fig 8C) and the upstream control element (UCE), which is a promoter element of poxviral early genes (Fig 8A). The UCE motif is a 15-nt long AT-rich element with any nucleotide at the 5<sup>th</sup> position followed by TGA [48]. This perfectly matched the extended UCS motif. The median distances from the 3' ends of the UCEs to the annotated TSSs of 84 Vaccinia virus ORFs displayed a median distance of 12 nt (Fig 8B) [48]. We annotated the TSSs of all pGKL genes based on our previous 5' RACE-PCR experiments (Materials and Methods, and S4 Table). The distances from the extended UCSs to the annotated TSSs of 15 pGKL-encoded ORFs (Fig 8D) had a median distance of 11 nt. Moreover, we found an adenosine residue to be the TSS nucleotide in all pGKL-encoded ORFs (S4 Table), similar to the TSSs of the poxviral early genes where purines are the dominant TSS bases [48].

To conclude, it appears that promoters of poxviral early genes and VLE genes are similar both with respect to their sequence and their spacing to the TSSs, implying a common origin.



**Fig 8. Upstream promoter elements of yeast VLEs show high sequence similarity to promoters of poxviral early genes.** (A) Vaccinia virus early promoter consensus motif termed upstream control element (UCE) calculated from 84 sequences identified from genome-wide RNA-sequencing experiments in ref. [48]. (B) Graph showing the number of UCE sequences as a function of their distances to the transcription start sites (TSS) of 84 ORFs as annotated in ref. [48]. (C) Extended promoter consensus motif of pGKL elements preceding 15 ORFs. This motif contains the upstream conserved sequence (UCS) which is universal among yeast VLEs. (D) Graph showing the number of extended UCS sequences as a function of their distances to the TSS of 15 pGKL-encoded ORFs as determined in 5' RACE-PCR experiments. For promoters with a putative initiator region the first adenosine residue in the region was considered to be the TSS. For more information concerning promoter characterization see [Materials and methods](#).

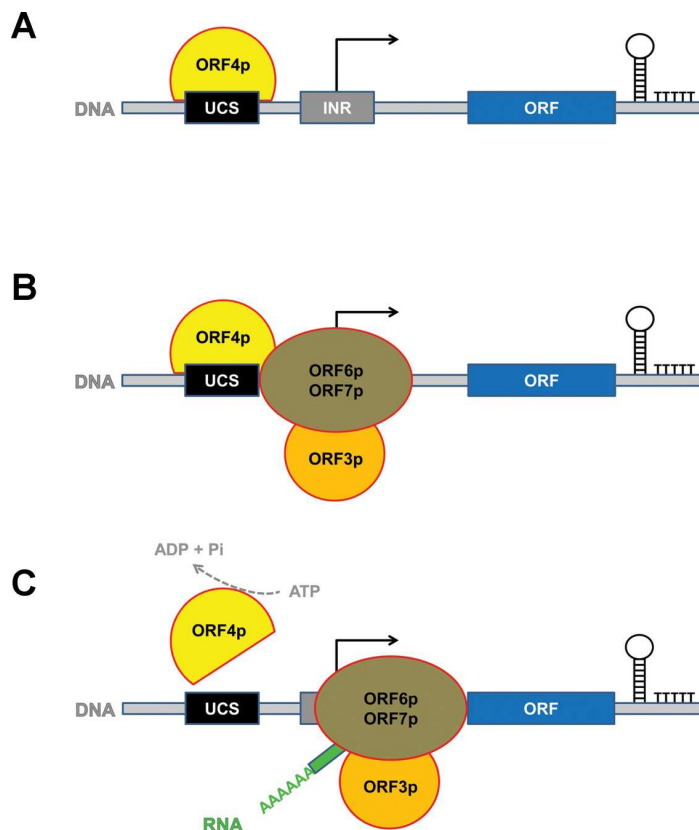
<https://doi.org/10.1371/journal.ppat.1007377.g008>

## Discussion

In this study we characterized the considerably underexplored transcription machinery of the yeast cytoplasmic linear double-stranded DNA virus-like elements. We used both experimental and bioinformatic approaches, and determined the composition and interactions of the transcription complex and presented a 3D model of its two main subunits. Further, we defined DNA sequences required for initiation and termination. For a model of the key aspects of transcription of the VLEs see Fig 9. Finally, our analyses provided evidence strongly suggesting that poxviruses and the yeast VLEs have a common origin.

## Composition of the VLE transcription machinery

Biochemical characterization of proteins encoded by yeast VLEs was shown to be challenging in the past. Expression of genes located on the pGKL elements seemed to be rather weak [15–17]. Also, it has been shown that expression of the K2ORF3p mRNA capping enzyme in routinely used *E. coli* systems was not possible, most likely due to the different codon usage



**Fig 9. Model of the yeast VLE transcription initiation and termination.** (A) Putative helicase K2ORF4p (ORF4p, yellow) binds to the VLE DNA, presumably to the upstream conserved sequence (UCS, black) which is related to the early promoter element of poxviruses. (B) K2ORF4p recruits the RNAP complex (ORF6p/ORF7p, brown) to the transcription initiation site, which usually contains the initiator region (INR, grey) responsible for RNAP slippage and subsequent 5' mRNA polyadenylation. (C) ATP hydrolysis by K2ORF4p releases it from the transcription preinitiation complex to allow RNAP to escape from the initiation site and produce mRNA (RNA, green) containing a 5' end poly (A) leader. This RNA can be subsequently 5' capped by the K2ORF3p viral-like mRNA capping enzyme (ORF3p, orange). Transcription termination most likely proceeds in a factor-independent manner that involves intrinsic terminators consisting of RNA stem loop structure(s) and 3' terminal U-tract.

<https://doi.org/10.1371/journal.ppat.1007377.g009>

dictated by the high AT content of VLE genes [32]. Recently, it was shown that yeast nuclear expression of VLE genes was impaired because the high AT content of VLE genes led to RNA fragmentation [33]. Accordingly, we failed to express recombinant K2ORF6p and K2ORF7p RNAP subunits in *E. coli*, *S. cerevisiae* and *K. lactis* expression systems.

To overcome this difficulty, we prepared modified and double-modified pGKL elements expressing the transcription components with tags. Using co-immunoprecipitations followed by mass spectrometry and Western blotting we demonstrated that *in vivo* the transcription machinery core complex consisted of the two RNAP subunits (K2ORF6p, K2ORF7p) and the mRNA capping enzyme (K2ORF3p). This interaction was independent of the presence of nucleic acids. Subsequently, we showed that the putative helicase (K2ORF4p) associated less tightly with both the VLE large RNAP subunit and the mRNA capping enzyme. This suggested the following molecular model of the VLE transcription components interactions. First, we propose that K2ORF6p and K2ORF7p interact directly because aa residues of  $\beta'$ a15 and  $\beta'$ a16, which are known to participate in intramolecular (*Bacteria*, *Eukarya*, *Archaea*) or intermolecular (*Archaea*) bonds in multisubunit RNAPs, are present also in K2ORF6p and K2ORF7p as revealed by our *in silico* analysis. Second, we suggest that the mRNA capping enzyme interacts directly with the RNAP complex although it is yet to be determined whether it is with K2ORF6p and/or K2ORF7p. Cellular mRNA capping enzymes are known to interact with the C-terminal domain of RNAP II  $\beta'$  homolog subunit [49]. However, a homologous C-terminal domain is present neither in Vaccinia virus nor in VLE RNAPs. Nevertheless, in Vaccinia virus, the heterodimeric VTF mRNA capping enzyme interacts directly with the RNAP complex, and it is thought to be present both during transcription initiation and elongation [50, 51]. By analogy, K2ORF3p may utilize a similar mode of interaction with its RNAP. Third, we propose that K2ORF4p binds to the RNAP core complex less tightly than the core subunits do between themselves. Vaccinia virus helicases VETF and NPH-I are known to interact with RNAP through an RNAP-associated protein of 94 kDa (RAP94), and this is specific for RNAP packaged in the virion [52–54]. VETF can interact with RNAP only in the presence of RAP94 [53]. However, there seem to be no RAP94 homologs outside poxviruses, which implies yet another undescribed mechanism of D6/D11 homologs binding to RNAP in other NCLDVs. This might apply also to K2ORF4p of the yeast VLEs. Future studies will have to address the exact mode of K2ORF4p binding to the transcription complex. However, due to relatively weak interaction of the putative helicase with the core of the VLE-specific transcription complex *in vivo*, we assume that it possibly acts as a dissociable transcription factor.

Finally, we found K2ORF4p, K2ORF3p and K2ORF6p to associate with pGKL-specific DNA *in vivo*. We propose that the transcription complex interacts directly with pGKL-specific DNA through RNAP subunits, and that the association of the mRNA capping enzyme with pGKL-specific DNA is indirect. Based on Vaccinia virus early transcription system analogy, we believe that the putative helicase also interacts directly with pGKL-specific DNA with the UCS promoter elements being the K2ORF4p binding sites. However, due to compact genomic organisation of pGKL elements precise mapping of the *in vivo* DNA binding sites of K2ORF4p using ChIP-seq would most likely turn out unsuccessful because resolution of the method would not match close spacing of the UCS elements. Our analysis indeed suggested close spacing of K2ORF4p binding sites *in vivo* because any pGKL region chosen for PCR amplification showed up to be enriched in anti-HA-K2ORF4p ChIP sample. Importantly, it seems that the transcription machinery of the yeast VLEs is remarkably if not entirely self-sufficient, because we did not find any cellular proteins to be specifically associated with the VLE large RNAP subunit using mass spectrometry analysis.

## Transcription initiation

Our previous 5' RACE-PCR experiments revealed short poly(A) leaders at the 5' mRNA ends of most pGKL-encoded genes [36]. These 5' poly(A) leaders were heterogeneous in length among individual transcripts (1–21 adenosines per molecule) and not complementary to the template DNA. Non-template 5' poly(A) leaders are a characteristic feature of Vaccinia virus intermediate and late gene mRNAs that occur due to slippage of RNAP at the INR of the promoter [37–40]. Moreover, these non-template 5' poly(A) leaders were also reported for some Vaccinia virus early genes [48, 55, 56]. Thus, it seems that it is a general feature of Vaccinia virus RNAP to slip on consecutive T residues in the INR element notwithstanding the promoter class [48].

It has been shown that 5' untranslated regions composed of 5' poly(A) leader sequences prior to start codon have a regulatory role in translation initiation [57, 58]. Using bioinformatics, we detected a putative INR element in promoters of pGKL genes whose transcripts were 5' polyadenylated. Using 5' RACE-PCR and mutagenesis of the putative INR, we confirmed that the 5' poly(A) leader of VLE transcripts was associated with the identified INR element *in vivo*, and it was generated by the same mechanism as shown for Vaccinia virus postreplicative transcripts [39]. In Vaccinia virus, however, mutations in the INR of postreplicative promoters completely abrogate marker gene expression [39, 59]. This was clearly not the case for INR of pGKL promoters because expression of *G418<sup>R</sup>* was used for selection of clones containing recombinant VLEs. To assess possible role of INR alterations on VLE gene expression, we prepared a strain (IFO1267\_pRKL1-9) where we introduced TACCC mutations to the TAAAC INR of toxin subunit gene *K1ORF4*, and this strain showed reduced killer toxin production (S8 Fig).

## Transcription termination

We mapped the 3' mRNA ends of all pGKL-genes using 3' RACE-PCR experiments. We identified 1–4 putative RNA stem loops close to the 3' mRNA terminus of each ORF. Although the putative stem loops displayed relatively high values of Gibbs free energy (median of −7.5 kcal/mol), these values were comparable to the genome-wide predicted intrinsic terminators in *Mycoplasma hyopneumoniae* (median of −8.0 kcal/mol), an organism with a similarly high AT content [60]. Further, RNA stem loops of bacterial intrinsic transcription terminators are usually followed by the typical 7–8 nt U-tract that promotes RNAP pausing at weak dA-rU DNA-RNA hybrid [61–63]. Interestingly, we detected T nucleotide enrichment in terminal 8 nt of VLE-specific 3' cDNA ends that corresponds to putative U-tract (S9 Fig). Importantly, we revealed a direct link between the putative RNA stem loop structure and the transcription termination pattern *in vivo*, suggesting an intrinsic transcription termination model for the yeast VLEs, similar to that in bacteria (reviewed in [64]). Even though we did not analyse the termination efficiency, we assume that transcription reads through at least some of the putative terminator sequences. Otherwise, functional expression of approximately half of the ORFs would not be possible due to the compact genomic organization inherent to pGKL elements. Future experiments will be required to understand this mechanism in more detail.

Even though Vaccinia virus RNAP, presumably related to VLE-specific RNAP, terminates transcription of early genes in a factor-dependent manner it was recently shown that RNA stem loops can influence both efficiency and location of transcription termination *in vitro* [65]. Therefore, proposed substantial reduction of VLE-specific RNAP ancestor might have contributed to adaptation of this enzyme to transcription termination induced by RNA stem loops and possible loss of auxiliary factors required for this process. In Vaccinia virus, the mRNA capping enzyme and NPH-I helicase are the auxiliary factors essential for transcription termination of early genes [26, 31]. Due to similarities of these viral proteins with K2ORF3p



and K2ORF4p [25, 28], respectively, the role of VLE mRNA capping enzyme and VLE putative helicase in transcription termination cannot be ruled out because of their interaction with both VLE-specific RNAP and DNA *in vivo*. However, we were not able to locate a specific termination signal (UUUUUNU) of Vaccinia virus early mRNAs that is required for RNAP pausing induced by viral mRNA capping enzyme [65] in VLE mRNAs. This suggests that unlike in poxviral early genes, VLE mRNA capping enzyme possibly does not play a direct role in VLE transcription termination which may instead solely use RNA stem loops for RNAP pausing. Nevertheless, possible roles of K2ORF3p, K2ORF4p, and/or other proteins in transcription termination are yet to be determined.

### ***In silico* 3D model of VLE RNAP**

Bioinformatic analysis of VLE-specific RNAP proved to be challenging due to its unique reduced architecture and great evolutionary distance from other multisubunit RNAPs. From the 3D model it is evident that VLE RNAP significantly differs from canonical RNAPs in several aspects:

- (i) Almost the entire clamp structure element is absent. Only a basal portion of the clamp formed by  $\beta$ a15,  $\beta$ a16 and  $\beta'$ a20 conserved regions is maintained. The clamp is a mobile RNAP element and its closure is important for high stability and processivity of the enzyme. The clamp conformation is regulated by interaction of universally conserved elongation factors NusG and Spt4/5 with the clamp coiled-coil motif [66]—an element likely missing in VLE RNAP. Therefore, it is highly unlikely that VLE transcription machinery could use cellular Spt4/5 to increase processivity.
- (ii) The lid and rudder elements are likely missing. The lid acts as a wedge to facilitate dislocation of RNA from the DNA-RNA hybrid molecule, and thereby maintains a constant size of the DNA-RNA hybrid between 7 to 10 base pairs [67]. However, it was shown, that lid-less bacterial and archaeal RNAPs were negatively affected when transcribing from ssDNA but not so much from dsDNA templates [68, 69]. By analogy, RNA displacement from DNA-RNA hybrid molecule should not be affected by absence of the lid in VLE RNAP transcribing double-stranded templates. The lid was also suggested to participate in bacterial intrinsic termination using stem loops [67]. However, bacterial RNAP without the lid was capable of intrinsic termination *in vitro* [69]. Therefore, we hypothesise that this structural feature is not crucial for intrinsic termination by VLE RNAP *in vivo*. The rudder element interacts with the upstream edge of the DNA-RNA hybrid [67]. Experiments using bacterial RNAP with a deleted rudder reported defects in transcription initiation and less stable elongation complexes [70]. This may correlate with the VLE-specific termination of transcription.
- (iii) The secondary-channel rim helices are missing. These helices are the binding sites for some transcription factors of multisubunit RNAPs, such as the transcription elongation factor TFIIS [71]. Therefore, it is highly unlikely that VLE transcription machinery could use cellular TFIIS to overcome pause sites and increase proofreading.

### **The evolutionary origin of the VLE transcription machinery**

A viral origin of the VLE genes encoding the mRNA capping enzyme and the putative helicase was suggested previously [1]. However, the same origin for RNAP genes was not expected because previous hypotheses proposed that those genes originated from ancestral yeast RNAP genes [1, 23] or that they were ancient representatives of multisubunit RNAP diversification due to their simplified architecture [45]. However, no phylogenetic analysis was conducted to



support the aforementioned hypotheses. Our results indicate that VLE RNAP has its evolutionary origin close to poxviruses. Poxviral RNAP, such as that of Vaccinia virus, lacks obvious  $\alpha$  subunit homologs [72], similarly to VLE RNAP, and is more simplified than eukaryotic RNAPs. Moreover, the constrained tree topology with enforced monophyly of VLE RNAPs with both eukaryotic RNAP I and RNAP II was rejected at a statistically significant level based on Expected Likelihood Weight test. This also correlates with the reported phylogenetic analysis of mRNA capping enzyme guanylyltransferase core regions [1]. Therefore, a reduction of poxviral RNAP instead of yeast RNAP to give rise to the VLE RNAP seems most plausible.

Our promoter analysis suggests that not just the VLE RNAP, but also the VLE promoters are related to nucleocytoplasmic large DNA viruses. We noticed sequence similarity between the UCE motif of Vaccinia virus early genes and the extended UCS motif of pGKL elements, as well as their similar location prior to the TSSs. Invariant G residue and several A residues in the minor groove were proposed to be the UCE nucleotides contacted by Vaccinia virus VETF helicase [27]. Presence of the invariant G residue and AT residues at other positions of the extended UCS motif suggests that K2ORF4p might contact UCS of the pGKL elements in a VETF-like fashion and act as an ATP-dependent transcription initiation factor. A sequence motif similar to UCE has been also identified in promoters of early genes of NCLDV of *Mimiviridae* family, although its location prior to the TSSs seems to differ [73].

A recent evolutionary hypothesis suggests that both yeast VLEs and nucleocytoplasmic large DNA viruses branched from a Polintovirus—eukaryotic dsDNA virus that had acquired transcription machinery genes required for escape from nucleus to cytoplasm [74]. Yeast VLEs and NCLDV then may have evolved separately [74]. It is difficult to dispute this hypothesis, also considering that transcription machinery genes similar to those encoded by the yeast VLEs and NCLDV have not been reported yet in the so far characterized Polintoviruses. Based on all our results, we propose that yeast VLEs might have evolved directly from a proto-poxvirus rather than from a common ancestor of VLEs and NCLDV.

In conclusion, the transcription apparatus of the yeast VLEs has most likely an origin close to poxviruses and uses transcription initiation mechanisms similar to those used by poxviral genes. Unlike poxviruses, however, VLEs are beneficial for the cell, and this exemplifies the ability of the cell to domesticate potentially pathogenic elements.

## Materials and methods

### Strains, plasmids and growth conditions

All of the strains used in this study are listed in Table 2. *Escherichia coli* cells were grown at 37°C in 2xTY medium which was supplemented with kanamycin (50 µg/ml) or ampicillin (100 µg/ml) for selection of transformants. Transformations of *E. coli* cells were performed by electroporation using Gene Pulser Xcell (Bio-Rad). *K. lactis* cells were grown at 28°C in YPD medium which was supplemented with G418 (250 µg/ml) and/or hygromycin B (200 µg/ml) for selection of transformants. Transformations of *K. lactis* cells were performed using the one-step LiCl method [75] and followed by five-hour incubation in non-selective conditions immediately after transformation. For detailed descriptions of plasmids and elements used in this study see S1 Table. Constructed pGKL elements were verified by PCR and subsequent sequencing of amplified products.

The nucleotide sequences of the primers used for construction, verification and sequencing of recombinant pGKL elements, and RACE-PCR amplification are listed in S2 Table. All polymerase chain reactions (PCRs) were performed using Taq DNA polymerase (Roche). PCRs for construction of recombinant pGKL elements were performed using mixture of Taq DNA polymerase (Roche) and Pwo DNA polymerase (Roche) in a 99:1 volume ratio, respectively.

Table 2. List of strains used in this study.

Strain	Genotype	Reference / Source
<b><i>E. coli</i> strains</b>		
XL-1 Blue	<i>endA1 gyrA96(nal<sup>R</sup>) thi-1 recA1 relA1 supE44 lac [F' Tn10(tet<sup>R</sup>) proAB lacI<sup>q</sup>Δ(lacZ)M15] hsdR17(r<sub>K</sub><sup>-</sup> m<sub>K</sub><sup>+</sup>)</i>	Stratagene
<b><i>S. cerevisiae</i> strains</b>		
S6/1	MATα	[76]
<b><i>K. lactis</i> strains</b>		
IFO1267	wild-type MATa pGKL1 <sup>+</sup> pGKL2 <sup>+</sup>	Institute for Fermentation, Osaka
IFO1267_pRKL1-1	pRKL1-1 <sup>+</sup> (K1UCR2-G418 <sup>R</sup> , orf2Δ) pGKL2 <sup>+</sup>	This study
IFO1267_pRKL1-2	pRKL1-2 <sup>+</sup> (K1UCR2*-G418 <sup>R</sup> , orf2Δ) pGKL2 <sup>+</sup>	This study
IFO1267_pRKL1-3	pRKL1-3 <sup>+</sup> (K1UCR2** -G418 <sup>R</sup> , orf2Δ) pGKL2 <sup>+</sup>	This study
IFO1267_pRKL1-4/2-4	pRKL1-4 <sup>+</sup> (ORF4-HA, HygB <sup>R</sup> ) pRKL2-4 <sup>+</sup> (yEGFP3-ORF6, G418 <sup>R</sup> )	This study
IFO1267_pRKL1-5	pRKL1-5 <sup>+</sup> (G418 <sup>R</sup> -K2ORF5 3' UTR, orf2Δ) pGKL2 <sup>+</sup>	This study
IFO1267_pRKL1-6	pRKL1-6 <sup>+</sup> (G418 <sup>R</sup> -K2ORF5 3' UTR <sup>o</sup> , orf2Δ) pGKL2 <sup>+</sup>	This study
IFO1267_pRKL1-7	pRKL1-7 <sup>+</sup> (G418 <sup>R</sup> -K2ORF5 3' UTR <sup>o</sup> , orf2Δ) pGKL2 <sup>+</sup>	This study
IFO1267_pRKL1-9	pRKL1-9 <sup>+</sup> (K1UCR4*-ORF4, HygB <sup>R</sup> ) pGKL2 <sup>+</sup>	This study
IFO1267_pRKL2-3	pGKL1 <sup>+</sup> pRKL2-3 <sup>+</sup> (FLAG-ORF6, G418 <sup>R</sup> )	This study
IFO1267_pRKL2-4	pGKL1 <sup>+</sup> pRKL2-4 <sup>+</sup> (yEGFP3-ORF6, G418 <sup>R</sup> )	This study
IFO1267_pRKL2-5	pGKL1 <sup>+</sup> pRKL2-5 <sup>+</sup> (yEGFP3-ORF6, ORF7-FLAG, G418 <sup>R</sup> , HygB <sup>R</sup> )	This study
IFO1267_pRKL2-6	pGKL1 <sup>+</sup> pRKL2-6 <sup>+</sup> (yEGFP3-ORF6, ORF3-HA, G418 <sup>R</sup> , HygB <sup>R</sup> )	This study
IFO1267_pRKL2-7	pGKL1 <sup>+</sup> pRKL2-7 <sup>+</sup> (yEGFP3-ORF6, HA-ORF4, G418 <sup>R</sup> , HygB <sup>R</sup> )	This study
IFO1267_pRKL2-8	pGKL1 <sup>+</sup> pRKL2-8 <sup>+</sup> (yEGFP3-ORF4, G418 <sup>R</sup> )	This study
IFO1267_pRKL2-9	pGKL1 <sup>+</sup> pRKL2-9 <sup>+</sup> (yEGFP3-ORF4, HA-ORF6, G418 <sup>R</sup> , HygB <sup>R</sup> )	This study
IFO1267_pRKL2-10	pGKL1 <sup>+</sup> pRKL2-10 <sup>+</sup> (yEGFP3-ORF4, ORF3-HA, G418 <sup>R</sup> , HygB <sup>R</sup> )	This study
IFO1267_pRKL2-11	pRKL2-11 <sup>+</sup> (ORF3-yEGFP3, G418 <sup>R</sup> )	This study
IFO1267_pRKL2-12	pGKL1 <sup>+</sup> pRKL2-12 <sup>+</sup> (HA-ORF6, HygB <sup>R</sup> )	This study
IFO1267_pRKL2-13	pGKL1 <sup>+</sup> pRKL2-13 <sup>+</sup> (HA-ORF4, HygB <sup>R</sup> )	This study
IFO1267_pRKL2-14	pGKL1 <sup>+</sup> pRKL2-14 <sup>+</sup> (ORF3-HA, HygB <sup>R</sup> )	This study
IFO1267_pRKL2-15	pGKL1 <sup>+</sup> pRKL2-15 <sup>+</sup> (ORF7-FLAG, HygB <sup>R</sup> )	This study

Detailed description of all plasmids/elements used in this study is listed in [S1 Table](#). UCR is the sequence extending from the AUG initiation codon up to and including the UCS of the selected ORF.

\*UCR sequence bearing one point mutation in putative initiator region (INR).

\*\*UCR sequence bearing two point mutations in putative initiator region (INR).

<sup>o</sup>3' UTR of K2ORF5 gene bearing mutations in putative Stem loop 2.

<sup>oo</sup>3' UTR of K2ORF5 gene bearing rescue mutations in putative Stem loop 2.

<https://doi.org/10.1371/journal.ppat.1007377.t002>

## Modification of pGKL elements using homology recombination *in vivo*

*K. lactis* IFO1267 strain was transformed with PCR-generated fragment consisting of 5' and 3' ends homologous to the part of the pGKL element to be modified and non-homologous part that introduced purification and/or detection tag (yEGFP3, HA-tag, FLAG-tag) into VLE-specific ORF together with a gene encoding resistance marker (G418 or hygromycin B) whose expression is driven by pGKL1-derived upstream control region (UCR, the sequence extending from the AUG initiation codon up to and including the UCS of the selected ORF). This type of construct was prepared by PCR or fusion PCR methodology (for details see [S3 Table](#)).

After PCR amplification and gel electrophoresis, corresponding fragments were purified using FavorPrep GEL/PCR Purification Kit (Favorgen) and used for transformation. After transformation, yeast cells were plated onto selective media and analyzed using agarose

electrophoresis for the presence of the modified pGKL element. Usually, it was possible to detect both, modified and wild-type target element directly after transformation. Colonies containing both (modified and wild-type variants of the target pGKL element) were selected and cultivated under selective conditions for approximately 60 generations and analyzed again using agarose electrophoresis. For subsequent analysis and preparation of double-modified pGKL elements, colonies containing only modified variant of the target pGKL element were used. Absence of unmodified variant of the respective pGKL element was also verified by PCR and subsequent electrophoresis.

### Isolation of pGKL elements and electrophoresis

Modified protocol based on [77] was used for analysis of the pGKL VLEs. Briefly, cells were grown on YPD plates supplemented with G418 and/or hygromycin B for three days, transferred into a microplate well and dried for 2 hours at 45°C. After complete drying the cells were resuspended in 40 µl of freshly prepared TESP buffer (20 mM Tris-HCl, pH 8, 50 mM EDTA, 2% SDS, 0.5 mg/ml pronase E) and dried overnight at 37°C. The next day the sample was completely resuspended in 40 µl of 1x DNA loading buffer (Fermentas). 15 µl of the sample was analyzed by electrophoresis (0.5% agarose, voltage 1 V/cm) for at least 20 hours. Subsequently, the gel was incubated in a solution containing ethidium bromide (0.5 µg/ml) and RNase A (50 µg/ml) for at least 3 hours, and then briefly incubated in distilled water.

### Co-immunoprecipitation and mass spectrometry

For the co-immunoprecipitation experiments, 100 ml of the yeast cells from the late exponential growth phase ( $OD_{600} = 4-8$ ) were used. The cells were harvested with centrifugation, washed with distilled water and lysed in 2–3 ml of ice-cold GFP-wash buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5 mM EDTA) supplemented with 1 mM PMSF, cOmplete Mini protease inhibitors (Roche) and 0.45 mm glass beads using Mixer Mill MM 301 (Retsch) at a frequency of 30/s for 5 min. The glass beads and cell debris were pelleted with centrifugation at 8 000 g for 5 min at 4°C. The lysates were clarified with centrifugation at 20 000 g for 20 min at 4°C. Co-immunoprecipitations were performed using 25 µl of GFP-Trap\_A (Chromotek) beads with gentle mixing overnight at 4°C. Mock immunoprecipitations were performed using 25 µl of empty agarose beads (Chromotek) with gentle mixing overnight at 4°C. Bound proteins were washed three times in 1 ml of ice-cold GFP-wash buffer with gentle mixing for 5 min at 4°C. The immunoprecipitated complexes were dissolved in 60 µl of 2X sample loading buffer (0.1 M Tris-HCl, pH 6.8, 20% glycerol, 2% β-mercaptoethanol, 4% SDS and 0.04% bromophenol blue), incubated for 5 min at 95°C and subjected to SDS-PAGE. The gel was stained with Coomassie Brilliant Blue G-250 or silver. Gel lanes or bands of interest were excised from the gel, digested with trypsin [78], and analysed by mass spectrometry. Identity of all detected proteins was also confirmed by MS/MS analysis. To test, whether certain co-immunoprecipitation is dependent on nucleic acids, the bound washed proteins were treated with 25U of Benzonase Nuclease (Novagen) in GFP-wash buffer for 30 min at 33°C and washed three times again, prior to SDS-PAGE analysis. Fraction of the bound washed proteins was taken and nucleic acids were eluted in 30 µl of TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA), incubated for 5 min at 95°C, and 1 µl of the nucleic acids was used for detection of DNA using PCR amplification for 22–25 cycles and subsequent electrophoresis.

### Chromatin immunoprecipitation

For the chromatin immunoprecipitation experiments, 50 ml of the yeast cells from the late exponential growth phase ( $OD_{600} = 4-8$ ) were used. The cells were cross-linked with

formaldehyde (1% final concentration) added to the growing culture for 40 min at 28°C. The cells were harvested with centrifugation, washed with distilled water and lysed in 2 ml of ice-cold non-denaturing lysis buffer (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 5 mM EDTA, 1% Triton X-100, 0.02% sodium azide) supplemented with 1 mM PMSF, cOmplete Mini protease inhibitors (Roche) and 0.45 mm glass beads using Mixer Mill MM 301 (Retsch) at a frequency of 30/s for 5 min. The lysate was sonicated (Qsonica Ultrasonic Processor Q700, 50% amplitude) sixty times with 10 sec pulses. The glass beads and cell debris were pelleted with centrifugation at 8 000 g for 5 min at 4°C. The lysates were clarified with centrifugation at 20 000 g for 20 min at 4°C and 50 µl aliquot of the clarified lysates representing chromatin before immunoprecipitation (Input) was taken. The DNA fragments were immunoprecipitated using 30 µl of mouse monoclonal anti-HA HA-7 agarose (Sigma Aldrich) beads or 20 µl of GFP-Trap\_A (Chromotek) beads with gentle mixing overnight at 4°C. The beads were then washed once in 1 ml of ice-cold Wash buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS) supplemented with single stranded salmon sperm DNA (100 µg/ml; Roche), twice in 1 ml of ice-cold Wash buffer, and once in 1 ml of ice-cold Final wash buffer (Wash buffer containing 500 mM NaCl), each time with gentle mixing for 5 min at 4°C. Immunocomplexes were then eluted from the beads in 120 µl of Elution buffer (1% SDS, 100 mM NaHCO<sub>3</sub>) for 30 min at 37°C (anti-HA HA-7 beads) or for 30 min at 65°C (GFP-Trap\_A beads). Eluted immunocomplexes and 50 µl of the clarified lysates representing chromatin before immunoprecipitation (Input) were mixed with 400 µl of TBS (50 mM Tris-HCl, pH 7.5, 150 mM NaCl) supplemented with 5 µl of proteinase K (20 mg/ml; Sigma Aldrich), and the cross-linking was reversed by incubation for 5 hr at 65°C. The immunoprecipitated and input DNA was isolated by phenol-chloroform extraction followed by ethanol precipitation supplemented with 1 µl of linear polyacrylamide (25 mg/ml; Sigma Aldrich), and then used for PCR amplification for 25–30 cycles followed by electrophoresis. PCR amplifications were carried out on 1/30 of the chromatin immunoprecipitation (ChIP) and 1/1200 of the chromatin before immunoprecipitation (Input) using primers listed in [S2 Table](#).

## Western blotting

SDS-PAGE gels were electroblotted onto Immun-Blot PVDF Membrane (Bio-Rad). The membranes were blocked in 5% non-fat dry milk (Hero) in a TBS-Tween buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl and 0.5% Tween-20) and incubated with a primary antibody overnight at 4°C. After washing in TBS-Tween buffer and blocking with 5% non-fat milk, the membranes were incubated with a goat anti-mouse HRP-conjugated antibody (1:5 000 dilution; Santa Cruz Biotechnology). Finally, after washing in TBS-Tween buffer, the membranes were immersed in a luminol detection solution and the signal was detected using ImageQuant LAS 4000 (GE Healthcare). To confirm the expression of the target protein and the successful immunoprecipitation, a mouse monoclonal anti-FLAG M2 antibody (1:1 000; Sigma Aldrich), mouse monoclonal anti-HA 6E2 antibody (1:1 000; Cell Signaling), and mouse monoclonal anti-GFP B-2 antibody (1:1 000; Santa Cruz Biotechnology) was used.

## RNA isolation, electrophoresis, reverse transcription, 5' and 3' RACE-PCR

25 ml of the yeast cells from the exponential growth phase ( $OD_{600} = 0.5–1$ ) were quickly pelleted and frozen. Total yeast RNA was isolated by the hot acidic phenol procedure followed by ethanol precipitation [79]. Remaining DNA was removed by DNA-free Kit (Ambion). The quality of RNA was assessed by electrophoresis and UV spectrophotometry [80].

In the case of 5' RACE, subsequent reverse transcription was carried out as follows: 1 µg of total yeast RNA and 0.15 µg of random hexamer primers (Invitrogen) were used for cDNA

synthesis using 100U of SuperScript III Reverse Transcriptase (Invitrogen). The cDNA was purified using High Pure PCR Product Purification Kit (Roche) and used for cDNA tailing using 800U of rTdT (Fermentas) and 0.5 mM dGTP (Roche) in 50  $\mu$ l reaction for 30 min at 37°C with subsequent heat inactivation of rTdT for 10 min at 70°C. For PCR amplification of cDNA ends, 2.5  $\mu$ l of the reaction mixture was used with olig2(dC)anchor primer and appropriate gene-specific primer for 35 cycles.

In the case of 3' RACE, 1  $\mu$ g of total yeast RNA was polycytidynylated using Poly(A) Tailing Kit (Applied Biosystems) and 2 mM CTP (Thermo Scientific) for 90 min at 37°C. Following reverse transcription was performed using 100U of SuperScript III Reverse Transcriptase (Invitrogen) and 1  $\mu$ g of oligo(dG)anch2 primer. The cDNA was purified using High Pure PCR Product Purification Kit (Roche) and 2.5  $\mu$ l of the purified cDNA was used for PCR amplification of cDNA ends with anch2 primer and appropriate gene-specific primer for 35 cycles.

In both RACE experiments, after PCR amplification and electrophoresis, obtained fragments were verified using restriction digestion and fragments exhibiting correct digestion pattern were gel-purified using FavorPrep GEL/PCR Purification Kit (Favorgen) and cloned to the pCR4-TOPO vector (Invitrogen). Vectors were transformed into *E. coli* XL-1 Blue cells (Stratagene), isolated using GenBond Plasmid FlexSpin Kit (Renogen Biolab) and sequenced using universal T7 promoter primer or T3 primer.

## Structure modelling

Modelling of the tertiary structure of the non-canonical RNA polymerase of the pGKL elements was carried out on the Robetta server [81]. Selection of the optimal template for comparative modelling by Robetta is automatic and five structural models for each sequence are given as the output. All predicted models were compared according to their tertiary structure with a database of known structures using DaliLite v. 3 [82]. Models that gave the best Z-scores to the known structures of canonical RNA polymerases were used for further analyses. In order to prevent partial misalignment of the sequences with the templates used for modelling, sequences of the K2ORF6p (P05472.1) and K2ORF7p (P05473.1) had to be split into parts that were submitted separately for the modelling.

Amino acid residues 1–693 of K2ORF6p were modelled by  $\beta$  subunit homolog of *Schizosaccharomyces pombe* RNAP II (PDB ID: 3H0G.B), residues 754–882 were modelled by  $\beta'$  subunit of *Thermus thermophilus* RNAP (PDB ID: 2A6H.D), and residues 894–974 were modelled by  $\beta'$  subunit homolog of *S. pombe* RNAP II (PDB ID: 3H0G.A). Amino acid residues 1–52 of K2ORF7p were modelled by  $\beta'$  subunit homolog of *S. pombe* RNAP II (PDB ID: 3H0G.A), and residues 103–132 were modelled by  $\beta'$  subunit homolog of *Saccharomyces cerevisiae* RNAP II (PDB ID: 1TWF.A).

Models of each part of K2ORF6p and K2ORF7p with the best Z-scores from previous comparison using DaliLite v. 3 were individually superimposed to *S. cerevisiae* RNA polymerase II elongation complex (PDB ID: 2NVQ) using DaliLite v. 3 in pairwise option and visualized together using PyMOL 1.3 software (The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC.). The template modeling score (TM-score) value [83], which measures the structural similarity, was calculated for each model:structure superimposition using TM-score (<http://zhanglab.ccmb.med.umich.edu/TM-score/>). Methionine 9 residue of the K2ORF6p sequence was considered to be the methionine 1 residue because it is thought that the second ATG within the ORF is the true start codon [6]. Therefore, numbering of K2ORF6p residues goes accordingly in this work.



## Phylogenetic analysis

Phylogenetic analysis was performed using  $\beta'$  subunit conserved sequence regions of multisubunit RNAPs that are present in all canonical RNAPs together with corresponding regions that were found in yeast VLE RNAPs. A single amino acid (aa) sequence was created by joining the aa sequences of K2ORF6p and K2ORF7p for each of the following organisms: *Kluyveromyces lactis* (P05472.1 and P05473.1), *Saccharomyces kluyveri* (CAA38625.1 and CAA38626.1), *Pichia acaciae* (CAJ57280.1 and CAJ57281.1) and *Pichia etchellsii* (CAC08226.1 and CAC08227.1). The joined sequences of RNAP subunits were aligned using default settings of Clustal Omega [84].  $\beta'$  subunit parts together with amino acid residues that did not correspond to previously [23, 24] identified  $\beta'$  subunit conserved regions were removed. Alignments of VLE RNAP  $\beta'$  subunit conserved regions were combined manually with previously published alignments [46] of  $\beta'$  subunit conserved regions of eukaryotic (120 sequences), archaeal (33 sequences), bacterial (40 sequences), plastid (4 sequences) and viral (34 sequences) RNAPs. Sequences of *Lausannevirus* (YP\_004346983.1 and YP\_004346982.1), *Marseillevirus marseillevirus* (YP\_003406803.1 and YP\_003406800.1), *Pandoravirus dulcis* (YP\_008318947.1), and *Pandoravirus salinus* (YP\_008436862.1) RNAP  $\beta'$  subunits were aligned using default settings of Clustal Omega, and their  $\beta'$  conserved regions were then manually added to the final alignment. PhyML 3.0 with Smart Model Selection [85] was used for initial maximum likelihood phylogenetic tree construction and selection of a substitution matrix and a model for rates across sites. For detailed phylogenetic analysis, unconstrained and alternative (constrained) topology maximum likelihood trees were constructed using IQ-TREE 1.6.1 [86] with the substitution model (LG+I+G4) found to be the best in the first-round analysis. Tree topologies were compared with IQ-TREE 1.6.1 using the Expected Likelihood Weight (ELW) test, and the approximately unbiased (AU) test. iTOL 3 [87] was used to view and analyze the phylogenetic trees.

## Promoter characterization

The known UCSs preceding all 15 pGKL-encoded genes [8] were used to construct manual alignments including sequences adjacent to the UCS at both upstream and downstream ends. WebLogo 2.8.2 (<http://weblogo.berkeley.edu/logo.cgi>) was used to create a consensus motif. The literature was searched for elements with similar motives known to be associated with promoter activity of multisubunit RNAPs.

Transcription start sites (TSSs) were annotated according to our previous 5' RACE-PCR experiments covering all pGKL genes [36]. Briefly, sequencing of individual cDNA clones of *K2ORF2*, *K2ORF3* and *K2ORF8* transcripts displayed a single peak TSS pattern, where more than 60% clones had the same initiator nucleotide and other initiator nucleotides were not represented in more than 25% of the clones. Sequencing of cDNA clones corresponding to transcripts of other pGKL-encoded ORFs displayed a multiple peaks TSS pattern, due to presence of a non-templated 5' poly(A) leader of a heterogeneous length. For those ORFs the first adenosine residue of the clones that could be aligned to the promoter template sequence was considered to be the initiator nucleotide.

## Assay of killer toxin activity

Filter sterilized culture medium was tested for the presence of the killer toxin activity by the agar well diffusion assay using *S. cerevisiae* S6/1 as a sensitive strain. Approximately  $2 \times 10^5$  of sensitive yeast cells were plated onto YPD plates (1% yeast extract, 2% peptone, 2% glucose, 2% agar) for testing of pGKL1 killer toxin activity. Wells were made with an 8 mm diameter cork borer and 100  $\mu$ l of filter sterilized culture medium or 100  $\mu$ l of serial dilution of the filter sterilized culture medium was pipetted into well.



## Supporting information

**S1 Fig. Electrophoretograms of isolated pGKL elements from yeast strains used in this study.** For each electrophoretogram DNA mass markers are indicated on the left, and native or recombinant pGKL elements are indicated on the right side, respectively. Recombinant pGKL elements are marked in red. M1, DNA molecular mass marker (Lambda DNA/Eco130I (StyI) Marker, Fermentas). M2, DNA molecular mass marker (GeneRuler 1 kb DNA Ladder, Thermo Scientific). IFO1267—*K. lactis* strain with wild-type pGKL elements. (TIF)

**S2 Fig. Identification of proteins associated with the capping enzyme encoded by the yeast VLEs.** (A) IFO1267\_pRKL2-4 (containing yEGFP3-K2ORF6p) and IFO1267\_pRKL2-11 (K2ORF3p-yEGFP3) cells were grown to late exponential phase. The cells were lysed, yEGFP3-K2ORF6p and K2ORF3p-yEGFP3 were affinity-purified using GFP-Trap agarose beads. Bound proteins were digested with trypsin, and then analyzed by mass spectrometry. Also, an aliquot of the beads was taken and bound proteins were eluted and resolved by SDS-PAGE, and the gel was silver-stained. Proteins identified by mass spectrometry are indicated with arrows on the right side. Proteins identified by mass spectrometry in previous experiments described in Table 1 and Fig 1 are indicated with dashed arrows on the right side. Bands corresponding to small RNA polymerase subunit were not clearly visible, presumably due to their smaller mass and weaker staining. M, protein molecular mass marker (PageRuler Prestained Protein Ladder, Fermentas); the respective molecular mass values are indicated on the left side. (B) K2ORF3p-associated proteins identified by mass spectrometry. The proteins identified, their molecular weight (MW), unique coverage, and the number of peptide types from IFO1267\_pRKL2-11 strain is listed. (TIF)

**S3 Fig. Physical association of the putative helicase and capping enzyme with the large RNAP subunit of the yeast VLEs is not dependent on nucleic acids.** (A) yEGFP3-K2ORF6p was purified with GFP-Trap agarose beads from lysates of the IFO1267\_pRKL2-6 strain cells. After washing the bound immunoprecipitated proteins ( $\alpha$ -GFP IP), the beads were split into two parts which were mock-treated (Benz-) and treated (Benz+) with Benzonase Nuclease to digest DNA and RNA. Then, the beads were extensively washed again and the bound proteins were eluted and analysed by Western blotting using anti-GFP ( $\alpha$ -GFP) and anti-HA ( $\alpha$ -HA) antibodies. Also, an aliquot of the beads was taken and bound nucleic acids were eluted and analyzed for presence of pRKL2-6 VLE DNA by PCR amplification for 25 cycles and visualized by electrophoresis. M, DNA molecular mass marker (GeneRuler 100 bp Plus DNA Ladder, Fermentas); respective molecular mass values of two DNA fragments (500 bp and 1 000 bp) are indicated. (B) yEGFP3-K2ORF6p was purified with GFP-Trap agarose beads from lysates of IFO1267\_pRKL2-7 strain cells. After washing the bound immunoprecipitated proteins ( $\alpha$ -GFP IP), the beads were split into two parts that were mock-treated (Benz-) and treated (Benz+) with Benzonase Nuclease to digest DNA and RNA. Then, the beads were extensively washed again and the bound proteins were eluted and analyzed by Western blotting using anti-GFP ( $\alpha$ -GFP) and anti-HA ( $\alpha$ -HA) antibodies. Also, an aliquot of the beads was taken and bound nucleic acids were eluted and analyzed for presence of pRKL2-7 VLE DNA by PCR amplification for 22 cycles and electrophoresis. M, DNA molecular mass marker (GeneRuler 100 bp Plus DNA Ladder, Fermentas); respective molecular mass values of two DNA fragments (500 bp and 1 000 bp) are indicated. (TIF)

**S4 Fig. 3' ends of VLE-specific mRNAs are located close to putative RNA stem loop structures.** This figure represents results of 3' RACE-PCR analysis of individual mRNAs

corresponding to all genes encoded by the pGKL elements. Total RNA was isolated from wild-type IFO1267 strain, DNase treated and 3' polycytidinylated. Reverse transcription was carried out using oligo(dG)anch2 primer. Purified cDNA was used for 3' RACE-PCR using anch2 primer and gene-specific primers listed in [S2 Table](#). After PCR amplification and electrophoresis, the identity of the products was verified using restriction digestion and fragments exhibiting correct digestion pattern were gel-purified, cloned to the pCR4-TOPO vector and sequenced. The upper sequence on the right side corresponds to the template (plasmid) DNA and its position in the pGKL genome is annotated using *K. lactis* pGKL1 (X00762.1) and pGKL2 (X07776.1) sequences. Sequences situated below represent individual sequenced cDNA clones (only distal part of 3' untranslated region is shown). Cytosine residues at the 3' end of cDNA corresponding to the RNA tail are omitted in this representation for clarity. RNA secondary structures close to 3' ends of VLE-specific mRNAs were predicted using default settings of RNAstructure Server (<http://rna.urmc.rochester.edu/RNAstructureWeb/>) [42]. Predicted RNA stem loops are displayed as cDNA nucleotide letters in circles on the left side, and the values of Gibbs free energy ( $\Delta G$ ) in kcal/mol are displayed for each structure. Stem loop distances from gene stop codon are shown as numbers of nucleotides (nt). Final nucleotides of the experimentally determined 3' ends of cDNA are shown as colored letters enlarged proportionally to their occurrence (in percent) in the sequenced clones when the same final nucleotide was detected in at least two independent clones. Predicted RNA stem loop structures and sequenced cDNA clones are listed as follows: (A) *K1ORF1*, (B) *K1ORF2*, (C) *K1ORF3*, (D) *K1ORF4*, (E) *K2ORF1*, (F) *K2ORF2*, (G) *K2ORF3*, (H) *K2ORF4*, (I) *K2ORF5*, (J) *K2ORF6*, (K) *K2ORF7*, (L) *K2ORF8*, (M) *K2ORF9*, (N) *K2ORF10*, (O) *K2ORF11*. (PDF)

**S5 Fig. 3' ends formation of VLE-specific mRNAs is dependent on 3' untranslated region and not dependent on promoter and coding sequence of the gene.** This figure represents results of 3' RACE-PCR analysis of individual mRNAs corresponding to the G418 and hygromycin B resistance genes expressed under control of *K1UCR2* and *K1UCR3*, respectively. (A) Schematic representation of recombinant pGKL2 elements where the resistance marker genes are inserted prior to *K2ORF5* gene in the same transcriptional orientation. (B) Total RNA was isolated from IFO1267\_pRKL2-3 and IFO1267\_pRKL2-12 strains, DNase treated and 3' polycytidinylated. Reverse transcription was carried out in the presence (+RT) and absence (-RT) of reverse transcriptase using oligo(dG)anch2 primer. Purified cDNA was used for 3' RACE-PCR using anch2 primer and gene-specific primers listed in [S2 Table](#). After PCR amplification the samples were analyzed in 1.8% agarose gel stained by ethidium bromide. M, DNA molecular mass marker (GeneRuler 100 bp Plus DNA Ladder, Fermentas). The respective values are indicated on the left side. Specific products that were cloned to the pCR4-TOPO vector and used for sequencing are labelled with asterisks. Predicted RNA stem loop structures and sequenced cDNA clones for mRNA 3' ends of (C) G418 and (D) Hygromycin B resistance genes are depicted as in [S4 Fig](#). (TIF)

**S6 Fig. ORF6 proteins of yeast VLEs show sequence similarity to  $\beta$  subunit conserved regions of multisubunit RNAPs.** Conserved regions are named according to ref. [46]. In the resulting sequence alignment the identity (black shading) was highlighted where the same amino acid residue occurred in  $\geq 50\%$  of the sequences and sequence similarity (gray shading) was highlighted where amino acid residue with similar properties occurred in  $\geq 50\%$  of the sequences. The numbers in brackets indicate the number of amino acid residues that were not displayed in this comparison. Sequence identity (in %) of consecutively numbered sequences is depicted in a table below each alignment. Local reliability of sequence alignments was evaluated using the Transitive Consistency Score (TCS) web server [88] and apart from the second

half of  $\beta$ a6 conserved region alignment all alignments showed good or average local reliabilities. Following sequences (with their accession numbers) were used for the alignment: *Thermus aquaticus*  $\beta$  (CAB65465.2), *Escherichia coli*  $\beta$  (AAC76961.1) *Methanocaldococcus jannaschii* B'' (Q58444.1), *Methanocaldococcus jannaschii* B' (Q60181.1), *Saccharomyces cerevisiae* Rpa135 (CAA95050.1), *Saccharomyces cerevisiae* Rpb2 (NP\_014794.1), *Saccharomyces cerevisiae* Rpc128 (CAA99422.1), *Vaccinia virus* Rpo132 (AAQ93241.1), *Kluyveromyces lactis* ORF6-pGKL2 (P05472.1), *Saccharomyces kluyveri* ORF6-pSKL (CAA38625.1), *Pichia acaciae* ORF6-pPac-1 (CAJ57280.1), *Pichia etchellsii* ORF6-pPE1B (CAC08226.1). (DOCX)

**S7 Fig. Results of statistical analysis of constrained trees for the  $\beta'$  subunit conserved regions of multisubunit RNAPs.** The trees are ranked from best to worst based on their likelihood. c-ELW, Expected Likelihood Weight. p-AU, p-value of approximately unbiased (AU) test. logL, log-likelihood. Plus signs denote the 95% confidence sets. Minus signs denote significant exclusion. All tests were performed with 10 000 resamplings using the REL method. (XLSX)

**S8 Fig. Mutations in the initiator region of VLE-specific promoter decrease gene expression.** This figure represents results of killer toxin activity assay. IFO1267\_pRKL1-9 (TACCC INR) and IFO1267 (control) cells were cultivated in YPD medium at 24°C. Aliquots were taken at 0, 3, 6 and 12 hours, and the culture medium was filter-sterilized, diluted, and assayed for the presence of the killer toxin activity by an agar well diffusion test using a lawn of *S. cerevisiae* S6/1 sensitive strain cells grown on YPD plates at 24°C for 2 days. Result from post-cultivation medium taken at 12 hours is shown. (TIF)

**S9 Fig. 3' ends of VLE-specific mRNAs contain putative U-tails.** This figure represents results of sequence analysis of all cDNA clones whose sequences were detected in at least two independent clones obtained from the 3' RACE-PCR analysis of mRNAs corresponding to all genes encoded by the pGKL elements. Those cDNA sequences (58 non-redundant sequences) were manually aligned according to their transcription termination site (TSS), and frequencies of individual nucleotides were calculated for each position within the 50 nt upstream of the TSS using Seqool 3.1 (<http://www.biossc.de/seqool/download.html>) and plotted. The occurrence of the T residue in the 8 3' terminal nucleotides of cDNA ends (including TSS) is  $\geq 50\%$ . (DOCX)

**S1 Table. Plasmids and virus-like elements used in this study.**  
(DOCX)

**S2 Table. Primers used in this study.**  
(DOCX)

**S3 Table. Primers, templates, and resulting PCR products used for modifications of pGKL elements.**  
(DOCX)

**S4 Table. Transcription start site positions.**  
(DOCX)

## Acknowledgments

We thank Natálie Suchánková and Jitka Vojáčková (Charles University, Czech Republic) for technical assistance.

## Author Contributions

**Conceptualization:** Michal Sýkora, Martin Pospíšek, Václav Vopálenský.

**Data curation:** Michal Sýkora, Václav Vopálenský.

**Formal analysis:** Michal Sýkora, Martin Pospíšek, Libor Krásný, Václav Vopálenský.

**Funding acquisition:** Martin Pospíšek.

**Investigation:** Michal Sýkora, Josef Novák, Silvia Mrvová, Václav Vopálenský.

**Methodology:** Michal Sýkora, Martin Pospíšek, Václav Vopálenský.

**Project administration:** Martin Pospíšek, Václav Vopálenský.

**Resources:** Martin Pospíšek, Libor Krásný.

**Supervision:** Martin Pospíšek, Václav Vopálenský.

**Visualization:** Michal Sýkora.

**Writing – original draft:** Michal Sýkora.

**Writing – review & editing:** Michal Sýkora, Libor Krásný, Václav Vopálenský.

## References

1. Jeske S, Meinhardt F, Klassen R. Extranuclear Inheritance: Virus-Like DNA-Elements in Yeast. In: Esser K, Löttge U, Beyschlag W, Murata J, editors. Progress in Botany. 68: Springer, Berlin, Heidelberg; 2007. p. 98–129.
2. Gunge N, Murata K, Sakaguchi K. Transformation of *Saccharomyces cerevisiae* with linear DNA killer plasmids from *Kluyveromyces lactis*. J Bacteriol. 1982; 151(1):462–464. PMID: [7045080](#).
3. Kikuchi Y, Hirai K, Hishinuma F. The yeast linear DNA killer plasmids, pGKL1 and pGKL2, possess terminally attached proteins. Nucleic Acids Res. 1984; 12(14):5685–5692. <https://doi.org/10.1093/nar/12.14.5685> PMID: [6379603](#).
4. Stam JC, Kwakman J, Meijer M, Stuitje AR. Efficient isolation of the linear DNA killer plasmid of *Kluyveromyces lactis*: evidence for location and expression in the cytoplasm and characterization of their terminally bound proteins. Nucleic Acids Res. 1986; 14(17):6871–6884. <https://doi.org/10.1093/nar/14.17.6871> PMID: [3763395](#).
5. Sor F, Fukuhara H. Structure of a linear plasmid of the yeast *Kluyveromyces lactis*; Compact organization of the killer genome. Curr Genet. 1985; 9(2):147–155. <https://doi.org/10.1007/Bf00436963>
6. Tommasino M, Ricci S, Galeotti CL. Genome organization of the killer plasmid pGK12 from *Kluyveromyces lactis*. Nucleic Acids Res. 1988; 16(13):5863–5878. <https://doi.org/10.1093/nar/16.13.5863> PMID: [3041369](#).
7. Gunge N, Tamaru A, Ozawa F, Sakaguchi K. Isolation and characterization of linear deoxyribonucleic acid plasmids from *Kluyveromyces lactis* and the plasmid-associated killer character. J Bacteriol. 1981; 145(1):382–390. PMID: [6257636](#).
8. Stark MJ, Boyd A, Mileham AJ, Romanos MA. The plasmid-encoded killer system of *Kluyveromyces lactis*: a review. Yeast. 1990; 6(1):1–29. <https://doi.org/10.1002/yea.320060102> PMID: [2180235](#).
9. Satwika D, Klassen R, Meinhardt F. Anticodon nuclease encoding virus-like elements in yeast. Appl Microbiol Biotechnol. 2012; 96(2):345–356. <https://doi.org/10.1007/s00253-012-4349-9> PMID: [22899498](#).
10. Kempken F, Hermanns J, Osiewacz HD. Evolution of Linear Plasmids. J Mol Evol. 1992; 35(6):502–513. <https://doi.org/10.1007/Bf00160211> PMID: [1474604](#).
11. Gunge N, Yamane C. Incompatibility of linear DNA killer plasmids pGKL1 and pGKL2 from *Kluyveromyces lactis* with mitochondrial DNA from *Saccharomyces cerevisiae*. J Bacteriol. 1984; 159(2):533–539. PMID: [6378881](#).
12. Romanos MA, Boyd A. A transcriptional barrier to expression of cloned toxin genes of the linear plasmid k1 of *Kluyveromyces lactis*: evidence that native k1 has novel promoters. Nucleic Acids Res. 1988; 16(15):7333–7350. <https://doi.org/10.1093/nar/16.15.7333> PMID: [3412888](#).

13. Kämpfer J, Meinhardt F, Gunge N, Esser K. In vivo construction of linear vectors based on killer plasmids from *Kluyveromyces lactis*: selection of a nuclear gene results in attachment of telomeres. *Mol Cell Biol*. 1989; 9(9):3931–3937. <https://doi.org/10.1128/mcb.9.9.3931> PMID: 2779572.
14. Kämpfer J, Esser K, Gunge N, Meinhardt F. Heterologous gene expression on the linear DNA killer plasmid from *Kluyveromyces lactis*. *Curr Genet*. 1991; 19(2):109–118. <https://doi.org/10.1007/BF00326291> PMID: 2065362.
15. Schründer J, Meinhardt F. An extranuclear expression system for analysis of cytoplasmic promoters of yeast linear killer plasmids. *Plasmid*. 1995; 33(2):139–151. <https://doi.org/10.1006/plas.1995.1015> PMID: 7597108.
16. Schickel J, Helmig C, Meinhardt F. *Kluyveromyces lactis* killer system: analysis of cytoplasmic promoters of the linear plasmids. *Nucleic Acids Res*. 1996; 24(10):1879–1886. <https://doi.org/10.1093/nar/24.10.1879> PMID: 8657569.
17. Schründer J, Gunge N, Meinhardt F. Extranuclear expression of the bacterial xylose isomerase (xylA) and the UDP-glucose dehydrogenase (hasB) genes in yeast with *Kluyveromyces lactis* linear killer plasmids as vectors. *Curr Microbiol*. 1996; 33(5):323–330. <https://doi.org/10.1007/s002849900122> PMID: 8875914.
18. Klassen R, Tontsidou L, Larsen M, Meinhardt F. Genome organization of the linear cytoplasmic element pPE1B from *Pichia etchellsii*. *Yeast*. 2001; 18(10):953–961. <https://doi.org/10.1002/yea.751> PMID: 11447601.
19. Schaffrath R, Soond SM, Meacock PA. The DNA and RNA polymerase genes of yeast plasmid pGKL2 are essential loci for plasmid integrity and maintenance. *Microbiology*. 1995; 141:2591–2599. <https://doi.org/10.1099/13500872-141-10-2591> PMID: 7582020.
20. Schaffrath R, Meinhardt F, Meacock PA. *ORF7* of yeast plasmid pGKL2: analysis of gene expression in vivo. *Curr Genet*. 1997; 31(2):190–192. <https://doi.org/10.1007/s002940050195> PMID: 9021138.
21. Jeske S, Tiggemann M, Meinhardt F. Yeast autonomous linear plasmid pGKL2: *ORF9* is an actively transcribed essential gene with multiple transcription start points. *FEMS Microbiol Lett*. 2006; 255(2):321–327. <https://doi.org/10.1111/j.1574-6968.2005.00082.x> PMID: 16448513.
22. Wilson DW, Meacock PA. Extranuclear gene expression in yeast: evidence for a plasmid-encoded RNA polymerase of unique structure. *Nucleic Acids Res*. 1988; 16(16):8097–8112. PMID: 3138657.
23. Ruprich-Robert G, Thuriaux P. Non-canonical DNA transcription enzymes and the conservation of two-barrel RNA polymerases. *Nucleic Acids Res*. 2010; 38(14):4559–4569. <https://doi.org/10.1093/nar/gkq201> PMID: 20360047.
24. Schaffrath R, Soond SM, Meacock PA. Cytoplasmic gene expression in yeast: A plasmid-encoded transcription system in *Kluyveromyces lactis*. *Biochem Soc Trans*. 1995; 23(1):S128. PMID: 7758690.
25. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res*. 1989; 17(12):4713–4730. <https://doi.org/10.1093/nar/17.12.4713> PMID: 2546125.
26. Deng L, Shuman S. Vaccinia NPH-I, a DEXH-box ATPase, is the energy coupling factor for mRNA transcription termination. *Gene Dev*. 1998; 12(4):538–546. <https://doi.org/10.1101/Gad.12.4.538> PMID: 9472022.
27. Broyles SS, Li J, Moss B. Promoter DNA Contacts Made by the Vaccinia Virus Early Transcription Factor. *J Biol Chem*. 1991; 266(23):15539–15544. PMID: 1869571.
28. Larsen M, Gunge N, Meinhardt F. *Kluyveromyces lactis* killer plasmid pGKL2: Evidence for a viral-like capping enzyme encoded by *ORF3*. *Plasmid*. 1998; 40(3):243–246. <https://doi.org/10.1006/plas.1998.1367> PMID: 9806862.
29. Mao XD, Shuman S. Intrinsic RNA (guanine-7) methyltransferase activity of the vaccinia virus capping enzyme D1 subunit is stimulated by the D12 subunit. Identification of amino acid residues in the D1 protein required for subunit association and methyl group transfer. *J Biol Chem*. 1994; 269(39):24472–24479. PMID: 7929111.
30. Schwer B, Hausmann S, Schneider S, Shuman S. Poxvirus mRNA cap methyltransferase—Bypass of the requirement for the stimulatory subunit by mutations in the catalytic subunit and evidence for intersubunit allostery. *J Biol Chem*. 2006; 281(28):18953–18960. <https://doi.org/10.1074/jbc.M602867200> PMID: 16707499.
31. Luo Y, Mao XD, Deng L, Cong PJ, Shuman S. The D1 and D12 Subunits Are Both Essential for the Transcription Termination Factor Activity of Vaccinia Virus Capping Enzyme. *J Virol*. 1995; 69(6):3852–3856. PMID: 7745734.
32. Tiggemann M, Jeske S, Larsen M, Meinhardt F. *Kluyveromyces lactis* cytoplasmic plasmid pGKL2: heterologous expression of *Orf3p* and proof of guanylyltransferase and mRNA-triphosphatase activities. *Yeast*. 2001; 18(9):815–825. <https://doi.org/10.1002/yea.728> PMID: 11427964.



33. Kast A, Voges R, Schroth M, Schaffrath R, Klassen R, Meinhardt F. Autoselection of Cytoplasmic Yeast Virus Like Elements Encoding Toxin/Antitoxin Systems Involves a Nuclear Barrier for Immunity Gene Expression. *Plos Genet.* 2015; 11(5):e1005005. <https://doi.org/10.1371/journal.pgen.1005005> PMID: 25973601.
34. Cormack BP, Bertram G, Egerton M, Gow NAR, Falkow S, Brown AJP. Yeast-enhanced green fluorescent protein (yEGFP): A reporter of gene expression in *Candida albicans*. *Microbiology.* 1997; 143:303–311. <https://doi.org/10.1099/00221287-143-2-303> PMID: 9043107.
35. Tokunaga M, Kawamura A, Hishinuma F. Expression of pGKL killer 28K subunit in *Saccharomyces cerevisiae*: identification of 28K subunit as a killer protein. *Nucleic Acids Res.* 1989; 17(9):3435–3446. <https://doi.org/10.1093/nar/17.9.3435> PMID: 2657659.
36. Vopalensky V, Sykora M, Masek T, Pospisek M. Messenger RNAs transcribed from yeast linear cytoplasmic plasmids possess unconventional 5' and 3' UTRs and suggest a novel mechanism of translation. *bioRxiv.* 2018. <https://doi.org/10.1101/325316>
37. Bertholet C, Van Meir E, ten Heggeler-Bordier B, Wittek R. Vaccinia Virus Produces Late mRNAs by Discontinuous Synthesis. *Cell.* 1987; 50(2):153–162. [https://doi.org/10.1016/0092-8674\(87\)90211-X](https://doi.org/10.1016/0092-8674(87)90211-X) PMID: 3036368.
38. Schwer B, Visca P, Vos JC, Stunnenberg HG. Discontinuous Transcription or RNA Processing of Vaccinia Virus Late Messengers Results in a 5' Poly(A) Leader. *Cell.* 1987; 50(2):163–169. [https://doi.org/10.1016/0092-8674\(87\)90212-1](https://doi.org/10.1016/0092-8674(87)90212-1) PMID: 3594569.
39. Davison AJ, Moss B. Structure of Vaccinia Virus Late Promoters. *J Mol Biol.* 1989; 210(4):771–784. [https://doi.org/10.1016/0022-2836\(89\)90108-3](https://doi.org/10.1016/0022-2836(89)90108-3) PMID: 2515287.
40. Schwer B, Stunnenberg HG. Vaccinia virus late transcripts generated in vitro have a poly(A) head. *EMBO J.* 1988; 7(4):1183–1190. <https://doi.org/10.1002/j.1460-2075.1988.tb02929.x> PMID: 3402436.
41. Yang ZL, Martens CA, Bruno DP, Porcella SF, Moss B. Pervasive Initiation and 3'-End Formation of Poxvirus Postreplicative RNAs. *J Biol Chem.* 2012; 287(37):31050–31060. <https://doi.org/10.1074/jbc.M112.390054> PMID: 22829601.
42. Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics.* 2010; 11:129. <https://doi.org/10.1186/1471-2105-11-129> PMID: 20230624.
43. Kwapisz M, Beckouet F, Thuriaux P. Early evolution of eukaryotic DNA-dependent RNA polymerases. *Trends Genet.* 2008; 24(5):211–215. <https://doi.org/10.1016/j.tig.2008.02.002> PMID: 18384908.
44. Rohe M, Schr nder J, Tudzynski P, Meinhardt F. Phylogenetic-Relationships of Linear, Protein-Primed Replicating Genomes. *Curr Genet.* 1992; 21(2):173–176. <https://doi.org/10.1007/Bf00318478> PMID: 1568258.
45. Iyer LM, Aravind L. Insights from the architecture of the bacterial transcription apparatus. *J Struct Biol.* 2012; 179(3):299–319. <https://doi.org/10.1016/j.jsb.2011.12.013> PMID: 22210308.
46. Lane WJ, Darst SA. Molecular Evolution of Multisubunit RNA Polymerases: Sequence Analysis. *J Mol Biol.* 2010; 395(4):671–685. <https://doi.org/10.1016/j.jmb.2009.10.062> PMID: 19895820.
47. Yutin N, Koonin EV. Hidden evolutionary complexity of Nucleo-Cytoplasmic Large DNA viruses of eukaryotes. *Virology.* 2012; 9:161. <https://doi.org/10.1186/1743-422X-9-161> PMID: 22891861.
48. Yang Z, Bruno DP, Martens CA, Porcella SF, Moss B. Genome-Wide Analysis of the 5' and 3' Ends of Vaccinia Virus Early mRNAs Delineates Regulatory Sequences of Annotated and Anomalous Transcripts. *J Virol.* 2011; 85(12):5897–5909. <https://doi.org/10.1128/JVI.00428-11> PMID: 21490097.
49. McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, et al. 5'-capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Gene Dev.* 1997; 11(24):3306–3318. <https://doi.org/10.1126/gad.11.24.3306> PMID: 9407024.
50. Luo Y, Hagler J, Shuman S. Discrete Functional Stages of Vaccinia Virus Early Transcription during a Single Round of RNA Synthesis *in Vitro*. *J Biol Chem.* 1991; 266(20):13303–13310. PMID: 1712778.
51. Hagler J, Shuman S. A freeze-frame view of eukaryotic transcription during elongation and capping of nascent mRNA. *Science.* 1992; 255(5047):983–986. <https://doi.org/10.1126/science.1546295> PMID: 1546295.
52. Mohamed MR, Niles EG. Interaction between nucleoside triphosphate phosphohydrolase I and the H4L subunit of the viral RNA polymerase is required for vaccinia virus early gene transcript release. *J Biol Chem.* 2000; 275(33):25798–25804. <https://doi.org/10.1074/jbc.M002250200> PMID: 10833518.
53. Yang ZL, Moss B. Interaction of the Vaccinia Virus RNA Polymerase-Associated 94-Kilodalton Protein with the Early Transcription Factor. *J Virol.* 2009; 83(23):12018–12026. <https://doi.org/10.1128/JVI.01653-09> PMID: 19759131.
54. Ahn BY, Gershon PD, Moss B. RNA Polymerase-associated Protein Rap94 Confers Promoter Specificity for Initiating Transcription of Vaccinia Virus Early Stage Genes. *J Biol Chem.* 1994; 269(10):7552–7557. PMID: 8125977.



55. Ahn BY, Jones EV, Moss B. Identification of the Vaccinia Virus Gene Encoding an 18-Kilodalton Subunit of RNA Polymerase and Demonstration of a 5' Poly(A) Leader on Its Early Transcript. *J Virol.* 1990; 64(6):3019–3024. PMID: [2335825](#).
56. Ink BS, Pickup DJ. Vaccinia virus directs the synthesis of early mRNAs containing 5' poly(A) sequences. *Proc Natl Acad Sci U S A.* 1990; 87(4):1536–1540. <https://doi.org/10.1073/pnas.87.4.1536> PMID: [1968253](#).
57. Shirokikh NE, Spirin AS. Poly(A) leader of eukaryotic mRNA bypasses the dependence of translation on initiation factors. *Proc Natl Acad Sci U S A.* 2008; 105(31):10738–10743. <https://doi.org/10.1073/pnas.0804940105> PMID: [18658239](#).
58. Xia XH, MacKay V, Yao XQ, Wu JH, Miura F, Ito T, et al. Translation Initiation: A Regulatory Role for Poly(A) Tracts in Front of the AUG Codon in *Saccharomyces cerevisiae*. *Genetics.* 2011; 189(2):469–478. <https://doi.org/10.1534/genetics.111.132068> PMID: [21840854](#).
59. Baldick CJ Jr., Keck JG, Moss B. Mutational Analysis of the Core, Spacer, and Initiator Regions of Vaccinia Virus Intermediate-Class Promoters. *J Virol.* 1992; 66(8):4710–4719. PMID: [1629951](#).
60. Fritsch TE, Siqueira FM, Schrank IS. Intrinsic terminators in *Mycoplasma hyopneumoniae* transcription. *BMC Genomics.* 2015; 16:273. <https://doi.org/10.1186/s12864-015-1468-6> PMID: [25886614](#).
61. Martin FH, Tinoco I Jr. DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Res.* 1980; 8(10):2295–2299. <https://doi.org/10.1093/nar/8.10.2295> PMID: [6159577](#).
62. d'Aubenton Carafa Y, Brody E, Thermes C. Prediction of Rho-independent *Escherichia coli* Transcription Terminators. A Statistical Analysis of their RNA Stem-Loop Structures. *J Mol Biol.* 1990; 216(4):835–858. [https://doi.org/10.1016/S0022-2836\(99\)80005-9](https://doi.org/10.1016/S0022-2836(99)80005-9) PMID: [1702475](#).
63. Gusarov I, Nudler E. The Mechanism of Intrinsic Transcription Termination. *Mol Cell.* 1999; 3(4):495–504. [https://doi.org/10.1016/S1097-2765\(00\)80477-3](https://doi.org/10.1016/S1097-2765(00)80477-3) PMID: [10230402](#).
64. Ray-Soni A, Bellecourt MJ, Landick R. Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annu Rev Biochem.* 2016; 85:319–347. <https://doi.org/10.1146/annurev-biochem-060815-014844> PMID: [27023849](#).
65. Tate J, Gollnick P. The role of vaccinia termination factor and *cis*-acting elements in vaccinia virus early gene transcription termination. *Virology.* 2015; 485:179–188. <https://doi.org/10.1016/j.virol.2015.05.008> PMID: [26280468](#).
66. Hirtreiter A, Damsma GE, Cheung ACM, Klose D, Grohmann D, Vojnic E, et al. Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. *Nucleic Acids Res.* 2010; 38(12):4040–4051. <https://doi.org/10.1093/nar/gkq135> PMID: [20197319](#).
67. Vassilyev DG, Vassilyeva MN, Perederina A, Tahirov TH, Artsimovitch I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature.* 2007; 448(7150):157–162. <https://doi.org/10.1038/nature05932> PMID: [17581590](#).
68. Naji S, Bertero MG, Spitalny P, Cramer P, Thomm M. Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. *Nucleic Acids Res.* 2008; 36(2):676–687. <https://doi.org/10.1093/nar/gkm1086> PMID: [18073196](#).
69. Touloukhonov I, Landick R. The Role of the Lid Element in Transcription by *E. coli* RNA Polymerase. *J Mol Biol.* 2006; 361(4):644–658. <https://doi.org/10.1016/j.jmb.2006.06.071> PMID: [16876197](#).
70. Kuznedelov K, Korzheva N, Mustaev A, Severinov K. Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. *EMBO J.* 2002; 21(6):1369–1378. <https://doi.org/10.1093/emboj/21.6.1369> PMID: [11889042](#).
71. Kettenberger H, Armache KJ, Cramer P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIIS. *Mol Cell.* 2004; 16(6):955–965. <https://doi.org/10.1016/j.molcel.2004.11.040> PMID: [15610738](#).
72. Knutson BA, Broyles SS. Expansion of poxvirus RNA polymerase subunits sharing homology with corresponding subunits of RNA polymerase II. *Virus Genes.* 2008; 36(2):307–311. <https://doi.org/10.1007/s11262-008-0207-3> PMID: [18264749](#).
73. Suhre K, Audic S, Claverie JM. Mimivirus gene promoters exhibit an unprecedented conservation among all eukaryotes. *Proc Natl Acad Sci U S A.* 2005; 102(41):14689–14693. <https://doi.org/10.1073/pnas.0506465102> PMID: [16203998](#).
74. Krupovic M, Koonin EV. Polintons: a hotbed of eukaryotic virus, transposon and plasmid evolution. *Nat Rev Microbiol.* 2015; 13(2):105–115. <https://doi.org/10.1038/nrmicro3389> PMID: [25534808](#).
75. Gietz RD, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* 2002; 350:87–96. [https://doi.org/10.1016/S0076-6879\(02\)50957-5](https://doi.org/10.1016/S0076-6879(02)50957-5) PMID: [12073338](#).

76. Woods DR, Bevan EA. Studies on the Nature of Killer Factor Produced by *Saccharomyces cerevisiae*. J Gen Microbiol. 1968; 51:115–126. <https://doi.org/10.1099/00221287-51-1-115> PMID: [5653223](#).
77. Pospíšek M, Palková Z. Microisolation of yeast nucleic acids on the microtitre plate without using lytic enzymes. Nucleic Acids Res. 1991; 19(18):5083. <https://doi.org/10.1093/nar/19.18.5083> PMID: [1923782](#).
78. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat Protoc. 2006; 1(6):2856–2860. <https://doi.org/10.1038/nprot.2006.468> PMID: [17406544](#).
79. Lin RJ, Kim DH, Castanotto D, Westaway S, Rossi JJ. RNA preparation from yeast cells In: Krieg PA, editor. A Laboratory Guide to RNA: Isolation, Analysis, and Synthesis: Wiley-Liss, New York; 1996. p. 43–50.
80. Mašek T, Vopálenský V, Suchomelová P, Pospíšek M. Denaturing RNA electrophoresis in TAE agarose gels. Anal Biochem. 2005; 336(1):46–50. <https://doi.org/10.1016/j.ab.2004.09.010> PMID: [15582557](#).
81. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. 2004; 32:W526–W531. <https://doi.org/10.1093/nar/gkh468> PMID: [15215442](#).
82. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucleic Acids Res. 2010; 38:W545–W549. <https://doi.org/10.1093/nar/gkq366> PMID: [20457744](#)
83. Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality. Proteins. 2004; 57(4):702–710. <https://doi.org/10.1002/prot.20264> PMID: [15476259](#).
84. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li WZ, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011; 7:539. <https://doi.org/10.1038/msb.2011.75> PMID: [21988835](#).
85. Lefort V, Longueville JE, Gascuel O. SMS: Smart Model Selection in PhyML. Mol Biol Evol. 2017; 34(9):2422–2424. <https://doi.org/10.1093/molbev/msx149> PMID: [28472384](#).
86. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015; 32(1):268–274. <https://doi.org/10.1093/molbev/msu300> PMID: [25371430](#).
87. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016; 44(W1):W242–W245. <https://doi.org/10.1093/nar/gkw290> PMID: [27095192](#).
88. Chang JM, Di Tommaso P, Lefort V, Gascuel O, Notredame C. TCS: a web server for multiple sequence alignment evaluation and phylogenetic reconstruction. Nucleic Acids Res. 2015; 43(W1):W3–W6. <https://doi.org/10.1093/nar/gkv310> PMID: [25855806](#).

### 10.1.2 Messenger RNAs of the yeast virus-like elements contain non-templated 5' poly(A) leaders and their expression is independent of eIF4E and Pab1

Václav Vopálenský, Michal Sýkora, Tomáš Mašek a Martin Pospíšek

*Frontiers in Microbiology*, v revizích

PMID: -

DOI: -

IF<sub>2018/2019</sub>: 4,259

Doplňkové informace:

<http://review.frontiersin.org/Document/DownloadSupplementaryMaterial?articleId=465106&userId=718917&roleId=18>

# Messenger RNAs of the yeast virus-like elements contain non-templated 5' poly(A) leaders and their expression is independent of eIF4E and Pab1

Václav Vopálenský<sup>1\*</sup>, Michal Sýkora<sup>1</sup>, Tomáš Mašek<sup>1</sup>, Martin Pospíšek<sup>1\*</sup>

<sup>1</sup>Faculty of Science, Charles University, Czechia

**Submitted to Journal:**  
Frontiers in Microbiology

**Specialty Section:**  
Virology

**Article type:**  
Original Research Article

**Manuscript ID:**  
465106

**Received on:**  
09 Apr 2019

**Frontiers website link:**  
[www.frontiersin.org](https://www.frontiersin.org)

### ***Conflict of interest statement***

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

### ***Author contribution statement***

VV and MP devised the project and the main conceptual ideas. VV, MS, MP and TM designed and performed the experiments. VV and MS created figures. MP received funding. All authors analyzed the data, discussed the results and contributed to the final manuscript.

### ***Keywords***

pGKL, poly(A) leader, poxvirus, eIF4E, Pab1, LSm1, VLE, virus-like element, linear cytoplasmic plasmid

### ***Abstract***

Word count: 159

We employed virus-like elements pGKL1,2 from *Kluyveromyces lactis* as a model to investigate the previously neglected transcriptome of the broader group of the yeast cytoplasmic linear dsDNA virus-like elements (VLEs). We performed 5' and 3' RACE analysis of all the pGKL1,2 mRNAs and found them not 3' polyadenylated and containing frequently uncapped 5' poly(A) leaders that are not complementary to the VLEs genomic DNA. The degree of 5' capping and/or 5' mRNA polyadenylation is specific to each gene and is controlled by the corresponding promoter region. Expression of pGKL1,2 transcripts is independent of eIF4E and Pab1 and is enhanced in *lsm1Δ* and *pab1Δ* strains. We suggested a model of a primitive regulation of pGKL1,2 gene expression where the degree of 5' mRNA capping and 5' non-template polyadenylation, together with the presence of negative regulators as Pab1 and Lsm1, play an important role. Our data also support a hypothesis of a close relationship between the yeast linear VLEs and poxviruses.

### ***Contribution to the field***

Extrachromosomal hereditary elements such as organelles, viruses, and plasmids significantly influence fitness and survival of the cell. We present a novel and complete view on previously neglected transcriptome of the yeast virus-like elements (cytoplasmic linear plasmids), represented by pGKL1 and pGKL2 from the kefir yeast *Kluyveromyces lactis*. We discovered that pGKL1,2 mRNAs frequently contain non-templately added 5' poly(A) leaders and lack 5' m7G caps. Such mRNA ends are very unique and have never been described in other eukaryotes and their viruses previously. We show that length of the 5' poly(A) mRNA leaders inversely correlates with degree of 5' mRNA capping. The degree of 5' capping and/or 5' mRNA polyadenylation is specific to each gene and is controlled by the corresponding promoter region that causes plasmid RNA polymerase slippage. Translation of pGKL1,2 transcripts is independent of eIF4E and Pab1 translation factors, and is also enhanced in *lsm1Δ* and *pab1Δ* strains. We suggested a model of a primitive regulation of pGKL1,2 gene expression where the degree of 5' mRNA capping and 5' non-template polyadenylation and negative regulators as PAB1 and Lsm1 play an important role. These findings, together with other available data, suggest a close relationship between the yeast virus-like elements pGKL1,2 and poxviruses.

### ***Funding statement***

This work was supported by the Czech Science Foundation [grant No. GBP305/12/G034]; the Charles University institutional project [SVV-260426]; the ELIXIR CZ research infrastructure project MEYS [Grant No: LM2015047] including access to computing and storage facilities and by MiCoBion project funded from EU H2020 (No 810224).

### ***Ethics statements***

(Authors are required to state the ethical considerations of their study in the manuscript, including for cases where the study was exempt from ethical approval procedures)

*Does the study presented in the manuscript involve human or animal subjects:* No

***Data availability statement***

Generated Statement: All datasets generated for this study are included in the manuscript and the supplementary files.

In review



# Messenger RNAs of the yeast virus-like elements contain non-templated 5' poly(A) leaders and their expression is independent of eIF4E and Pab1

Václav Vopálenský\*, Michal Sýkora, Tomáš Mašek, Martin Pospíšek\*

Laboratory of RNA Biochemistry, Faculty of Science, Department of Genetics and Microbiology, Charles University, Prague, Czech Republic

## \* Correspondence:

Václav Vopálenský

vaclav.vopalensky@natur.cuni.cz

Martin Pospíšek

martin.pospisek@natur.cuni.cz

**Keywords:** virus-like element, VLE, linear cytoplasmic plasmid, pGKL, poly(A) leader, poxvirus, eIF4E, Pab1, Lsm1.

## Abstract

We employed virus-like elements pGKL1,2 from *Kluyveromyces lactis* as a model to investigate the previously neglected transcriptome of the broader group of the yeast cytoplasmic linear dsDNA virus-like elements (VLEs). We performed 5' and 3' RACE analysis of all the pGKL1,2 mRNAs and found them not 3' polyadenylated and containing frequently uncapped 5' poly(A) leaders that are not complementary to the VLEs genomic DNA. The degree of 5' capping and/or 5' mRNA polyadenylation is specific to each gene and is controlled by the corresponding promoter region. Expression of pGKL1,2 transcripts is independent of eIF4E and Pab1 and is enhanced in *lsm1Δ* and *pab1Δ* strains. We suggested a model of a primitive regulation of pGKL1,2 gene expression where the degree of 5' mRNA capping and 5' non-template polyadenylation, together with the presence of negative regulators as Pab1 and Lsm1, play an important role. Our data also support a hypothesis of a close relationship between the yeast linear VLEs and poxviruses.

## INTRODUCTION

Yeast double-stranded DNA virus-like elements (VLEs; also termed linear cytoplasmic plasmids) have been found in the cytoplasm of yeast species from nine genera, including *Kluyveromyces*, *Debaryomyces*, *Saccharomyces*, *Saccharomycopsis*, *Wingea*, and *Pichia* (Fukuhara, 1995; Gunge and Tokunaga, 2004). The overall genetic organization of all yeast VLEs is identical with that of the most

studied VLEs, pGKL1 (also termed K1) and pGKL2 (also termed K2), from the yeast *Kluyveromyces lactis*, where their presence is associated with the killer phenotype (Gunge et al., 1981). Killer strains contain 50 to 100 copies of each of the two linear VLEs per cell (Gunge, 1983). The pGKL1 and pGKL2 VLEs have compact and extremely AT-rich genomes sized 8874 and 13447 bp, respectively, and carry terminal inverted repeats with proteins covalently attached to their 5' ends (Hishinuma et al., 1984; Stam et al., 1986; Tommasino et al., 1988).

The smaller VLE, pGKL1, contains four open reading frames (ORFs): two of them (*K1ORF2* and *K1ORF4*) encode precursors of killer toxin subunits (Stark et al., 1984; Stark and Boyd, 1986; Tokunaga et al., 1989); *K1ORF3* is involved in an immunity phenotype in an unknown manner (Tokunaga et al., 1987), and *K1ORF1* codes for the pGKL1 VLE-specific DNA polymerase and a terminal protein (Fukuhara, 1987; Jung et al., 1987). Eleven ORFs have been reported in the larger pGKL2 VLE, which provides vital functions for pGKL1/pGKL2 maintenance in host cells (Schaffrath and Meacock, 1995). Some functions have been attributed to more than half of the proteins encoded by the ORFs carried by the pGKL2 VLE. *K2ORF2* codes for a pGKL2-specific DNA polymerase and a terminal protein (Tommasino et al., 1988; Takeda et al., 1996). *K2ORF3* codes for a virus-like mRNA capping enzyme (Larsen et al., 1998; Tiggemann et al., 2001). *K2ORF4* encodes a putative helicase, which is involved in transcription of VLE-specific mRNAs (Stark et al., 1990; Sýkora et al., 2018), *K2ORF5* codes for a single-stranded DNA-binding protein, and *K2ORF10* encodes a protein bound to the terminal inverted repeats of both pGKL VLEs. All these proteins, with the exception of K2Orf3p and K2Orf4p, are probably involved in pGKL1/2 replication (Stark et al., 1990; McNeel and Tamanoi, 1991; Tommasino, 1991; Schaffrath and Meacock, 2001; Jeske et al., 2007). *K2ORF6* and *K2ORF7* code for putative subunits of a pGKL1/2-specific RNA polymerase (Wilson and Meacock, 1988; Schaffrath et al., 1997). We showed in our parallel study that K2Orf6p, K2Orf7p and K2Orf3p interact and form a pGKL transcription machinery core complex, which loosely interacts with a putative helicase K2Orf4p (Sýkora et al., 2018). No function has been assigned to the four remaining ORFs (*K2ORF1*, 8, 9, and 11) (Figure 1). All ORFs are likely expressed independently. Each gene is preceded by an upstream conserved sequence (UCS) motif, which is located approximately -30 nucleotides from the putative start codon for pGKL1 UCSs (AT<sup>A</sup>/C<sup>T</sup>TGA) or up to -100 nucleotides in the case of pGKL2 UCSs (ATNTGA) (for review see (Stark et al., 1990; Gunge and Tokunaga, 2004)). Sequences located between an AUG initiation codon and its respective UCS (inclusive) are called UCR (Upstream Control Region) and act as

promoters in *sensu lato* (Meinhardt et al., 1994; Schrunder and Meinhardt, 1995; Schrunder et al., 1996).

In summary, the yeast linear pGKL VLEs represent a model and the most studied example of a unique world of the yeast cytoplasmically localized linear double-stranded DNA VLEs. They code for their own replication, transcription and an RNA-modification apparatus, making them remarkably independent of their host cells and evoking thoughts about their similarity to other cytoplasmically localized linear DNA genomes such as viruses belonging to the families *Poxviridae* or *Asfarviridae* (Jeske et al., 2007). In our parallel study we analyzed a transcription apparatus of the pGKL VLEs. We demonstrated that pGKL RNA polymerase K2Orf6p/K2Orf7p is related to viral RNA polymerases from the *Poxviridae* and *Iridoviridae* families and an extended pGKL UCS motif shows high similarity to the UCE motif of the vaccinia virus early promoters (Sýkora et al., 2018).

The mechanism of action of the *K. lactis* heterotrimeric killer toxin is well understood. A  $\gamma$ -subunit of the toxin inhibits the growth of sensitive yeast cells by cleaving their tRNAs at the 3' side of the modified wobble nucleoside 5-methoxycarbonylmethyl-2-thiouridine (mcm5s2U), which leads to the cell cycle arrest in G1 phase (Butler et al., 1991; Huang et al., 2005; Lu et al., 2005; Jablonowski et al., 2006; Meineke et al., 2012). Readers are referred to references by Kast *et al.* and Meineke *et al.* for more information concerning action of the *K. lactis* toxin and related ribotoxins (Meineke et al., 2012; Kast et al., 2014).

In this work, we focus on the yet uncharacterized transcriptome of the yeast cytoplasmic linear dsDNA VLEs and the possible mechanism of translation initiation used by the VLE-specific mRNAs. We analyzed transcripts from all pGKL ORFs by 5' RACE method. Our analyses revealed that mRNAs transcribed from pGKL VLEs contain poly(A) leaders of variable lengths at their 5' ends that are not complementary to the VLE genomic DNA sequence, and thus must be synthesized in a template-independent manner. We demonstrate that VLE promoters directly determine the formation and length of the mRNA 5' end leaders and their capping, pointing to a possible role of the VLE RNA polymerase in this unusual phenomenon. We also show that pGKL mRNAs are not 3' polyadenylated and that most of them even lack the m<sup>7</sup>G cap at their 5' ends. These findings directed our interest to the translation of pGKL mRNAs. Using *in vitro* and *in vivo* approaches, we demonstrate that translation of VLE-specific mRNAs is independent of the major cellular cap-binding protein, translation initiation factor 4E (eIF4E), and that translation of VLE-specific mRNAs is also independent of the Pab1 and Lsm1 proteins that were speculated previously to be possible

translational enhancers and/or stabilizers of cellular and viral mRNAs bearing 5' poly(A) stretches (Bergman et al., 2007; Gilbert et al., 2007). Thus, our results suggest the existence of an unusual translation initiation mechanism.

## MATERIALS AND METHODS

### Modification of pGKL VLEs using homologous recombination *in vivo*

In general, *K. lactis* IFO1267 cells were transformed with a PCR-generated fragment (for an example of such a construct, see Figure 5A) consisting of 5' and 3' ends homologous to the part of the pGKL VLE to be modified, a non-homologous part that introduces mutation/s into a VLE-specific protein coding sequence, and a gene encoding a resistance marker (typically against G418 or hygromycin B) whose transcription is driven by the UCR of *K1ORF1* or *K1ORF2*. This type of construct was prepared by fusion PCR as follows. The first primary PCR product was amplified from a native pGKL2 VLE using Pfu DNA polymerase (Fermentas) and the primers ORF3\_SAM\_del\_F1 and ORF3\_SAMdel\_R1. The second primary PCR product was amplified using Pfu DNA polymerase (Fermentas) and the primers KL\_orf6C\_Flag2F and ORF3\_SAMdel\_R2 (see Table S1) from a modified pGKL1 VLE whose *K1ORF2* coding sequence we had replaced with *G418* resistance gene previously. Both primary PCR products, which had defined overlapping ends, were synthesized as follows: 5 min at 95°C; then 30 cycles of 30 sec at 94°C, 30 sec at 56.2°C, and 90 sec at 72°C; and finally, 10 min at 72°C. Then, 1 µl of each PCR mixture was used without purification for a second PCR reaction using Pfu DNA polymerase (Fermentas), the primers ORF3\_SAM\_del\_F1 and ORF3\_SAMdel\_R2 and following conditions: 5 min at 95°C; then 35 cycles of 30 sec at 94°C, 30 sec at 60.1°C, and 2 min at 72°C; and finally, 10 min at 72°C). The resulting fusion PCR product was purified using agarose gel electrophoresis and transformed into *K. lactis* IFO1267. After five hours of cultivation under nonselective conditions, cells were plated onto solid medium containing G418 (250 µg/ml). DNA content of selected clones was analyzed for the presence of the modified pGKL VLE using agarose gel electrophoresis. Both modified and wild-type target VLEs were detected within a single clone after transformation. Colonies containing both the modified and wild-type target VLEs were selected and cultivated under selective conditions for approximately 60 generations and their DNA contents were analyzed using agarose gel electrophoresis. For subsequent analysis, only colonies containing a modified version of the target VLE were used. Site-specific integration of the PCR cassette into a pGKL2 VLE was evaluated by PCR with primers (in\_Kan\_rev1 and in\_ORF3\_forw) followed by sequencing of a gel-purified PCR fragment.

## RNA purification, electrophoresis, reverse transcription, and 5' and 3' RACE

Total yeast RNA was purified by the hot acidic phenol procedure (Lin et al., 1996). Remaining DNA was removed by a DNA-free Kit (Ambion) according to the manufacturer's protocol. The quality of RNA was assessed by electrophoresis according to the protocol by Masek et al. (Masek et al., 2005). For 5' RACE, RT-PCR was carried out as follows: 0.5 µg of total yeast RNA and 0.15 µg of random primers (Invitrogen) were used for cDNA synthesis using 100 U of SSC III Reverse Transcriptase (Invitrogen) in a 20 µl reaction (25°C for 10 min, 50°C for 99 min, 70°C for 15 min). After the cDNA purification using the High Pure PCR Product Purification Kit (Roche), 800 U of recombinant TdT (Roche) and 0.5 mM dGTP (Fermentas) were used for cDNA tailing in 1x TdT buffer without CoCl<sub>2</sub> for 30 min at 37°C with subsequent TdT inactivation at 70°C for 10 min. For amplification of cDNA, 2.5 µl of the reaction mixture was used for the following PCR with the universal olig2(dC) anchor primer and an appropriate gene-specific primer. In the case of 3' RACE, 0.5 µg of total yeast RNA was polycytidynlated using a Poly(A) Tailing Kit (Applied Biosystems) with 2 mM CTP in a total volume of 25 µl for 90 min at 37°C. Reverse transcription was then performed using oligo(dG)anch2 primer as indicated above. After cDNA purification using the High Pure PCR Product Purification Kit (Roche), 2.5 µl of the purified cDNA was used for a subsequent PCR reaction with the universal primer anch2 and an appropriate gene-specific primer. In both types of RACE experiments, following PCR amplification and electrophoresis, the corresponding fragments were purified from gel using a FastBack DNA minispin kit (Renogen Biolab), cloned into a pCR4-TOPO plasmid using the TOPO strategy and sequenced using the universal T7 promoter primer and/or a T3 primer. All primers used in this study are listed in Table S1.

## eIF4E expression and purification

The pGEX4T2::CDC33 plasmid was transformed into the expression strain *E. coli* BL21(DE3) (Merck). A 500 ml culture was grown at 28°C in 2xTY medium supplemented with ampicillin (100 µg/ml) until the OD<sub>600</sub> reached 0.20. At this point, the culture was cooled to 15°C, cultivated for 4 hours and then induced by 0.1 mM isopropyl β-D-thiogalactopyranoside. Production of recombinant protein was carried out with shaking at 15°C for another 14-16 hours. Lysates were prepared using the B-PER Protein Extraction Reagent (Thermo Scientific) according to the manufacturer's instruction and stored at -70°C for further use.

## mRNA/eIF4E binding assay

On the day of the binding experiment, bacterial lysate containing GST-eIF4E fusion protein was thawed in an ice bath, centrifuged for 10 min at 15000 g (4°C), and GST-eIF4E was purified using GST-affinity chromatography in a batch setup. In brief, 1 ml of the lysate was incubated for 80 min with 100 µl of glutathione-Sepharose 4 Fast Flow resin (GE Healthcare) at 4°C and washed four times with at least 40 volumes of ice-cold 1x phosphate-buffered saline (PBS). After the last washing step, the GST-eIF4E fusion protein bound to the glutathione-Sepharose was resuspended in 200 µl of buffer I (20 mM HEPES, pH 7.5; 0.1 mM EDTA, pH 8.0; 100 mM KCl; and 1 mM β-mercaptoethanol) and mixed with 3 µg of DNase I-treated total RNA purified from *K. lactis* IFO1267. Next, 2 µl of the reaction mixture were removed for subsequent analysis using RT-PCR and real-time quantitative PCR (qRT-PCR). The rest of the mixture was incubated for two hours at room temperature and washed six times (with approximately 70 volumes in the each washing step) with buffer I. After the last washing step, the GST-eIF4E-glutathione-Sepharose slurry containing bound mRNA was subjected to RT-PCR and detection of specific mRNAs. All buffers used for protein purification contained Complete (EDTA-free) Protease Inhibitor (Roche), and all buffers used for handling RNA contained RNasin Ribonuclease Inhibitor (Promega).

Two microliters of reverse transcriptase reaction (obtained from eIF4E-mRNA *in vitro* binding assay) were subjected to PCR amplification as follows (5 min at 94°C; then 35 cycles of 30 sec at 94°C; 30 sec at 50°C; 45 sec at 72°C; and finally, 4 min at 72°C) using Taq DNA polymerase (Roche) and *HGT1* and *K2ORF5* gene-specific primers (listed in Table S1).

As a control, total RNA used for the experiment was reverse transcribed and subjected to Real-Time PCR amplification (Figure S3).

## Transfer of cytoplasmic VLEs to yeast strains by incomplete mating

The *S. cerevisiae* strain YAT547 (*kar1*) containing pGKL VLEs (Gunge and Yamane, 1984) and isogenic *S. cerevisiae* ρ<sup>0</sup> strains bearing the wild-type *CDC33* gene (CWO4p<sup>0</sup>*CDC33*wt) and its temperature-sensitive mutations (CWO4p<sup>0</sup>*cdc33-1* and CWO4p<sup>0</sup>*cdc33-42*, for exact genotypes see Table S2) were cultured separately in YPD medium at 24°C overnight. Cells were pelleted the next day and resuspended in 200 µl of water, and 5 µl from each of the two required strains (approximately 3x10<sup>6</sup> cells) were mixed on the YPD agar plate and further incubated for 5 hours at 24°C. Cells were then scraped into liquid YPD medium, incubated at 24°C with gentle shaking overnight and plated on selective plates. The resulting strains were tested for auxotrophic markers, mating type and for killer toxin production under both permissive and non-permissive conditions.



Strains displaying a temperature-sensitive phenotype (when applicable), killer toxin production, growth on SD medium lacking methionine, no visible growth on SD medium lacking uracil and a MATa mating type were used for subsequent analyses.

#### **pGKL VLEs purification and electrophoresis**

For the analysis of pGKL VLEs, a protocol based on that of Pospisek and Palkova (Pospisek and Palkova, 1991) was used. Briefly, cells were grown for three days on a dish containing the appropriate antibiotics, transferred into a microplate well and dried for 2 hours at 45°C. After they had dried completely, cells were resuspended in 40 µl of freshly prepared TESP buffer (20 mM Tris-Cl, pH 8; 50 mM EDTA-NaOH, pH 8; 2 % SDS; and 0.5 mg/ml pronase E) and dried at 37°C overnight. The sample was completely resuspended the next day in 40 µl of 1x DNA loading buffer (Fermentas). Then, 15 µl of the sample were analyzed using agarose electrophoresis (0.5 % agarose in TAE buffer, 1 V/cm) for at least 20 hours. After the electrophoresis, the gel was incubated in a solution containing ethidium bromide (0.5 µg/ml) and RNase A (50 µg/ml) for 3 hours.

#### **Removal of mitochondrial DNA using ethidium bromide treatment**

Isogenic *S. cerevisiae* strains bearing the wild-type *CDC33* gene (CWO4*CDC33*wt) and its temperature-sensitive mutations (CWO4*cdc33-1* and CWO4*cdc33-42*, for exact genotypes see Table S2) (Altmann et al., 1989; Altmann and Trachsel, 1989) were cultivated for two days in synthetic drop-out minimal medium (SD) that lacked leucine and tryptophan (SD-TL) and contained ethidium bromide at a final concentration of 25 µg per ml. Afterwards, these cultures were diluted five thousand times, cultivated again for two days and then diluted and cultivated once more. Loss of mitochondria was verified by the lack of colony growth on medium with glycerol as the sole carbon source and by DAPI staining followed by fluorescence microscopy.

#### **Plasmid construction for killer toxin production in CWO4 yeast strains**

The *CDC33* (eIF4E) gene was amplified from *S. cerevisiae* genomic DNA using Pfu DNA polymerase (Fermentas) with eIF4Ef and eIF4Er primers containing NcoI and HindIII restriction sites as follows: 5 min at 95°C; then 25 cycles of 30 sec at 94°C, 30 sec at 55°C, and 2 min at 72°C; and finally, 10 min at 72°C. The PCR cassette was digested and inserted into the NcoI and HindIII restriction sites of pGEX-4T2 (GE Healthcare) to generate a pGEX4T2::CDC33 construct. A pYX212 yeast shuttle plasmid (Ingenius) was digested with EcoRI and XhoI restriction

endonucleases. A sequence coding for K1 toxin was excised from a pYX213::M1 (Valis et al., 2006) plasmid using the same restriction enzymes and ligated into the digested pYX212 plasmid to generate the plasmid pYX212::M1. All clones were verified by restriction endonuclease digestion and sequencing. All primers and plasmids used in this study are listed in Table S1 and Table S3, respectively.

### Transformation of yeast cells

All yeast transformations with plasmid DNA were performed with the one-step LiCl method (Gietz and Woods, 2002). The cells were grown in a shaker at 28°C in SD minimal medium lacking selected amino acids or nucleotide bases to ensure plasmid maintenance. Transformation using a PCR fragment for homologous recombination was carried out in the same way with the exception of including a five-hour incubation under nonselective conditions immediately after transformation prior to plating.

### Killer toxin production in modified CWO4 yeast strains

Tested strains (50 ml) were cultured in an appropriate medium at 24°C with shaking. When the  $A_{600}$  reached 0.9, cells were harvested and washed three times with the same medium preheated to 24°C. One-half of the culture (25 ml) was mixed with 25 ml of fresh medium preheated to 24°C and then cultivated at 24°C for more than 20 hours. The second half of the tested culture (25 ml) was mixed with 25 ml of fresh medium preheated to 52°C and then cultivated at 37°C for more than 20 hours. Aliquots (2 ml) were taken at 0, 3, 6 and 12 hours. Samples were harvested, sterilized using 0.2  $\mu$ m filters and assayed for the presence of the killer toxin activity in the culture medium by an agar well diffusion assay using *S. cerevisiae* S6/1 as a sensitive strain.

### Assay of killer toxin activity

Filter sterilized culture medium was tested for the presence of the killer toxin activity by the agar well diffusion assay using *S. cerevisiae* S6/1 as a sensitive strain. Approximately  $2 \times 10^5$  of sensitive yeast cells were plated onto OSS1 agar plates (5.3% worth agar; 0.7% agar; 1% glucose; 1M sorbitol; 0.002% methylene blue; buffered to pH 4.7 with sodium citrate) for testing of M1 killer toxin activity or onto YPD plates (1% yeast extract; 2% peptone; 2% glucose; 2% of agar) for testing of pGKL1 killer toxin activity. Wells were made with an 8 mm diameter cork borer and 100  $\mu$ l of filter sterilized culture medium or 100  $\mu$ l of serial dilution of the filter sterilized culture medium was pipetted into

well. For quantitative results, inhibition zones were measured using a calibrated digital microscope after 48 hours of incubation at 24°C.

#### ***PBP1*, *PAB1* and *LSM1* genes deletion.**

The *PBP1*, *PAB1* and *LSM1* genes were deleted from the chromosomes of the *K. lactis* IFO1267 strain using a loxP-*G418*-loxP cassette (Guldener et al., 1996). Briefly, the cassette for the deletion of the *PBP1* gene was amplified from the pUG6 plasmid by PCR with the primers KL\_pbp1-del\_Rev and KL\_pbp1-del\_For. The PCR product was resolved by agarose electrophoresis, purified, and transformed into *K. lactis* IFO1267. The deletion of the *PBP1* gene was verified by PCR. This modified strain was transformed with the pSH65 plasmid (Guldener et al., 2002), incubated for 5 hours in nonselective medium and plated on solid medium supplemented with phleomycin (400 µg/ml). Monocolonies were grown on YPD plates with phleomycin for two days and subsequently tested for their ability to grow on G418 (250 µg/ml) selective plates. The excision of the *G418* cassette was verified by PCR for selected colonies that did not grow in the presence of G418. These colonies were cultivated for five days with daily dilution under nonselective conditions (YPD medium only), resulting in the loss of the Cre-containing pSH65 plasmid. The resulting yeast strain, *K. lactis* IFO1267*pbp1*Δ, was used for subsequent *PAB1* deletion using the primers KL\_pab1-del\_For and KL\_pab1-del\_Rev in a similar fashion with the exception of the *G418* cassette excision. *K. lactis* IFO1267 was used for *LSM1* deletion using primers KL\_lsm1-del\_For and KL\_lsm1-del\_Rev in a similar fashion with the exception of *G418* cassette excision. The PCR reactions were carried out similarly (5 min at 95°C; then 30 cycles of 30 sec at 94°C, 30 sec at 55°C, and 4 min at 68°C; and finally, 10 min at 68°C). The nucleotide sequences of the primers used for the verification of the gene disruption cassettes are summarized in Table S1.

#### **Killer toxin production in *K. lactis* strains with deletion of *PBP1*, *PAB1* and *LSM1* genes**

Test cultures (10 ml) were incubated overnight in YPD medium supplemented with G418 (250 µg/ml) at 28 °C with shaking. In the next day, the optical density of tested strains was measured and strains were inoculated to an initial OD<sub>600</sub> of 0.035 into 50 ml of fresh medium preheated to 28°C, and then cultivated at 28°C for approximately 80 hours. Aliquots (2 ml) were taken during the experiment. Cells were harvested using centrifugation and the supernatant was sterilized using 0.2 µm filters. After serial two-fold dilution the supernatant samples were assayed for the presence of the killer toxin activity by a standard agar well diffusion assay using *S. cerevisiae* S6/1 as a sensitive strain and YPD agar plates (Figure S4A). Width of the inhibition zone was independently measured

at > 20 positions within a single zone using a calibrated digital microscope (Figure S4C) and the mean width of the inhibition zone normalized to the concentration of production cells (I) was calculated (Figure S4B).

#### Strains, plasmids, VLEs, primers and statistical analysis

Please follow Supplementary Material for the additional information concerning statistical analyses, strains, primers, plasmids and VLEs used.

## RESULTS AND DISCUSSION

### VLEs transcripts are uncapped and contain a non-template 5' poly(A) leader

We modified a method of rapid amplification of 5' cDNA ends (5' RACE) and used it for analyzes of the 5' ends of the pGKL transcripts. Briefly, we purified total RNA from the *K. lactis* IFO1267 strain, treated it with DNase I, synthesized cDNA using SuperScript III reverse transcriptase and random primers, G-tailed the cDNA using dGTP and terminal deoxynucleotidyl transferase (TdT) and amplified the 5' regions corresponding to transcripts from individual ORFs by PCR with an anchored oligo(dC) primer and an appropriate gene-specific primer (for primer sequences refer to Table S1). Reverse transcriptases, SuperScript III including, can overcome the 5'- 5' triphosphate bond between the 5' m<sup>7</sup>G cap and the first nucleotide of the nascent mRNA transcripts which enables detection whether the mRNA is capped or not (Davison and Moss, 1989a; Schwer et al., 1998; Schmidt and Mueller, 1999). PCR fragments were verified by digestion with restriction endonucleases and cloned into a pCR4-TOPO plasmid. From 20 to 79 independent clones corresponding to the mRNA 5' regions were sequenced and analyzed for each ORF. We used *K. lactis* *ACT* transcript (EnsemblFungi Id: KLLA0\_D05357g), which codes for actin and is transcribed by the RNA polymerase II from a nuclear gene as a control 5' capped mRNA.

The 5' RACE analysis of the pGKL transcripts revealed a unique composition of their 5' ends. With exception of some viral transcripts, eukaryotic mRNAs generally contain an m<sup>7</sup>G cap structure at their 5' ends, which is a prerequisite for their stability and efficient translation initiation (Ross, 1995; Mukherjee et al., 2012; Ramanathan et al., 2016). Even though the pGKL VLEs encode their own capping enzyme (*K2Orf3p*), we found only three ORFs (*K2ORF2*, *K2ORF3* and *K2ORF8*) of the fifteen encoded by the pGKL VLEs, of which at least 86% of the transcripts were 5' capped. Messenger RNAs encoded by other pGKL1/2 ORFs were 5' capped less frequently (Table S4). We haven't detected any capped transcripts for *K2ORF10* and found only 4% to 8% capped mRNAs

among *K1ORF2*, *K1ORF3*, *K2ORF7* and *K2ORF11* transcripts. Interestingly, *K2ORF10* and *K2ORF7* are essential for the pGKL1/2 VLE maintenance in the yeast cells (Schaffrath et al., 1997; 1999). Indispensability of *K2ORF11* for replication of the pGKL VLEs has not been tested yet. *K2ORF10* belongs to the highly expressed pGKL genes (McNeel and Tamanoi, 1991; Schaffrath et al., 1999). *K1ORF2* and *K1ORF3* code for the relatively highly secreted *K. lactis* killer toxin and a corresponding immunity, respectively (Stark and Boyd, 1986; Tokunaga et al., 1989). Absence and/or low occurrence of the guanosine cap at the 5' ends of some pGKL mRNAs thus apparently does not substantially affect expression of the corresponding genes.

Occurrence of 5' capping in transcripts from other pGKL1/2 ORFs lay between these two extremes within the interval of 17% to 55% of capped mRNAs (Table S4 and Table S5). Thereto, all pGKL1/2 genes gave rise to transcripts containing a short leader of non-template adenosine nucleotides (1-21) at their 5' ends that could not have been continuously transcribed from the VLEs genomes. Contrary to that, *K. lactis* actin mRNA revealed 5' guanosine caps and no 5' poly(A) leaders at 100% of the analyzed mRNA sequences (Figure 2A, Table S4 and Table S5). Presence of 5' guanosine cap in *K2ORF8* mRNAs was also confirmed by oligo-capping method (Maruyama and Sugano, 1994) (Figure S1).

We calculated the number of non-templated adenosine nucleotides in the 5' leaders of the mRNAs of each pGKL ORF as well as the minimal number of templated and non-templated adenosines found at the 5' ends of their uncapped transcripts and the fractions of capped mRNAs and mRNAs containing non-templated 5' adenosine nucleotides. We also calculated medians and means of the numbers of non-templated adenosine nucleotides added per mRNA molecule from each pGKL ORF. These data are summarized in Table S4 and Table S5A. Further analysis of the obtained data showed that occurrence of 5' capping is inversely correlated with the number of the non-templately added 5' adenosines (Figure 3 and Table S6). Longer 5' poly(A) leaders were found in mRNAs transcribed from ORFs that yielded less frequently capped or completely uncapped transcripts (Figure 3, Table S4 and Table S5A). Congruently, the 5' cap is more frequently present on the transcripts containing few or zero non-templated adenosine nucleotides at their 5' ends (Table S5A. *K2ORF2*, *K2ORF3*, *K2ORF4*, *K2ORF5*, *K2ORF8*). However, both non-templated 5' poly(A) leaders and mRNA caps were often found within the identical transcripts (Figure 2A and Table S5A).

More than half of all pGKL ORFs (9) contained 5' untranslated regions (UTRs) composed exclusively of adenosine nucleotides (Table S4 and Table S5A) regardless of their templated or non-

templated origin, thus resembling adenosine-rich yeast internal ribosome entry sites (IRESs) (Gilbert et al., 2007). The mRNAs transcribed from pGKL1 promoters tended to have longer and uncapped 5' poly(A) leaders than pGKL2 mRNAs (Figure 4, Table S4 and Table S5A). The 5' ends of VLE mRNAs resemble the organization of intermediate and late vaccinia virus mRNAs (Schwer et al., 1987; Ahn and Moss, 1989), late cowpox virus mRNAs (Patel and Pickup, 1987) and even some plant mRNAs (Gowda et al., 2006) that also contain non-templated poly(A) leaders at their 5' ends. However, all the vaccinia mRNA transcripts are assumed to be fully capped (Kates and Beeson, 1970; Wei and Moss, 1974; 1975; Boone and Moss, 1977). Considering the aforementioned examples and combining our findings with the knowledge that the domain organization of the K2Orf3p putative capping enzyme resembles that of the poxviral capping enzyme (Larsen et al., 1998) allow us to speculate on the similarity between pGKL VLEs and poxviruses, at least at the level of transcription and posttranscriptional modification. This hypothesis is further supported by our parallel work showing high similarity of pGKL1/2 RNA polymerase and promoters to poxviral RNA polymerases and promoters of early poxviral genes, respectively (Sýkora et al., 2018).

Our 5' RACE analyses also identified three new functional UCS sites and the corresponding alternative transcription start sites of *K1ORF4*, *K2ORF3* and *K2ORF4* (Table S5A, marked as blue boxes).

### The pGKL promoters determine 5' end mRNA capping and non-templated polyadenylation

Mapping of the 5' mRNA ends allowed us to determine putative UCS sites, UCR regions and transcription start sites (TSSs) for all the pGKL1 and pGKL2 ORFs. *K1ORF4*, *K2ORF3* and *K2ORF4* contain two putative UCS sites each. We evaluated nucleotide sequences of all these UCS sites and refined their previously reported consensus sequence, ATNTGA (Klassen and Meinhardt, 2007), by calculating the occurrences of all 4 nucleotides in the variable N position: A = 50%, G = 11%, C = 22%, and T = 17%. The TSS (+1 nucleotide) is usually located in a short window 17 to 22 nt downstream from the 3' end of the UCS; the median distance is 19 nt, and the extreme values are 13 and 25 nt. Transcription initiates at the adenosine nucleotide. Two exceptions were present in the minor transcripts of *K2ORF2* and *K2ORF4* which started with G and T, respectively. The minor upstream *K2ORF4* UCS ends with a guanosine nucleotide (ATATGG), an exception to the consensus sequence, and serves a non-canonical TSS, initiating with a T that is located only 13 nt downstream from the UCS and thus suggesting a slightly different mode of transcription initiation than in other UCRs.



The UTRs are very short regardless of the presence or absence of 5' non-template leaders. Exceptions to this rule are the *K2ORF2* UTR ( $\approx 60$  nt), *K2ORF3* <sub>long</sub>UTR ( $\approx 150$  nt), *K2ORF3* <sub>short</sub>UTR ( $\approx 60$  nt) and *K2ORF5* UTR ( $\approx 90$  nt) (Table S5A). Mapping the 5' ends of pGKL transcripts allowed us to classify genes encoded by pGKL VLEs into two major groups: genes transcribed to frequently 5' capped mRNAs with low average number of 5' non-templated adenosine nucleotides (*K2ORF2*, *K2ORF3*, *K2ORF4* and *K2ORF8*; > 50% of transcripts are 5' capped and median of added 5' non-template adenosines  $\approx 0$  per mRNA), and a larger group of genes producing transcripts less frequently 5' capped but with a higher degree of 5' non-templated polyadenylation (occurrence of 5' capped transcripts varies between 0% and 50% and median of added 5' non-templated adenosines lies between 2 and 7) (Table S4).

Our results are in a very good agreement with three previous studies which used various setup of classical primer extension method for 5' UTR mapping. Collectively, those studies identified uninterrupted clusters of major and minor TSSs for *K1ORF1*, *K1ORF2*, *K1ORF3*, *K1ORF4*, *K2ORF5* and *K2ORF9* genes (Romanos and Boyd, 1988; Schaffrath and Meacock, 1995; Jeske et al., 2006). Lengths of 5' UTRs detected by us and calculated as a sum of median lengths of 5' poly(A) leaders and templated parts of the corresponding 5' UTRs (Table S4) were in agreement with the distances of previously suggested TSS clusters from the respective AUG codons of all these genes (Romanos and Boyd, 1988; Schaffrath and Meacock, 1995) except *K2ORF9*, for which the median length of the 5' UTR detected by us was approximately 1 nucleotide longer (Jeske et al., 2006). Our results show that a remarkable heterogeneity in 5' UTR lengths of pGKL transcripts is caused by the heterogeneity in lengths of the non-templated 5' poly(A) leaders (Figure 4, Table S4 and S5A) rather than by a presence of multiple TSSs. The results also suggest that possible DNA polymerase slippage during PCR amplification and/or sequencing did not significantly alter determination of the 5' poly(A) leader length in our experiments.

We wanted to investigate whether UCR sequences can themselves contain signals that determine degree of the 5' end polyadenylation and frequency of 5' capping of the pGKL transcripts. For this purpose we decided to prepare a yeast strain with altered pGKL VLEs. Genomes of the pGKL VLEs are extremely tightly packed (Figures 1 and 5) (Stark et al., 1990) and can be manipulated only by homologous recombination *in vivo* (Kamper et al., 1989). For genetic engineering of pGKL VLEs we developed a unique fusion PCR strategy followed by homologous recombination *in vivo* that allowed us to precisely modify selected pGKL regions including introduction of point mutations. Using this approach we prepared a *K. lactis* IFO1267 strain bearing a

new cytoplasmic linear VLE pRKL2-1 (for details of the method, refer to Figure 5 and Materials and Methods section). The pRKL2-1 linear VLE is a pGKL2 derivative that contains a *G418<sup>R</sup>* resistance marker (*G418<sup>R</sup>*) under the control of *K1UCR2* and *K2ORF2* artificially controlled by *K1UCR1* (Figure 5). We found that the presence of a cap at the 5' end of pRKL2-1 mRNAs coding for *K2ORF2* and *G418<sup>R</sup>*, as well as the presence and length of a 5' poly(A) leader, were both directly influenced by the UCR sequence used for expression of these genes. Replacement of the wild-type UCR sequence of the *K2ORF2* gene by the *K1ORF1* UCR sequence (*K1UCR1*, Figure 5) led to a remarkable switch from nearly fully capped and almost not 5' polyadenylated transcripts (91.7% capped transcripts, average of added non-template adenosine nucleotides per mRNA molecule  $\approx 0$ ) to largely uncapped and 5' polyadenylated transcripts (7.7% capped transcripts, average of added non-template adenosine nucleotides per mRNA molecule  $\approx 5$ ). These values correspond very well to the level of 5' capping and non-template polyadenylation of the wild type *K1ORF1* mRNAs expressed from the pGKL1 VLE under control of natural *K1UCR1* (Figure 5, Figure S2 and Table S7). The same trend was also observed for the *G418<sup>R</sup>* transcripts, transcription of which was controlled by *K1UCR2*. The majority of these transcripts (96.4%) were uncapped, but all of them were 5' polyadenylated in exactly the same manner as the natural *K1ORF2* transcripts. The average number of non-templated adenosine residues in the 5' leader was comparable in both *G418<sup>R</sup>* and *K1ORF2* genes, controlled by *K1UCR2* (Figure 5, Figure S2 and Table S7). These findings strongly support our recent experiments showing that point mutations within the TSS regions of the pGKL genes can substantially influence 5' polyadenylation of the pGKL mRNAs (Sýkora et al., 2018) and further support the hypothesis that differences in the 5' end formation, both 5' capping and polyadenylation, of the yeast linear VLE mRNAs are controlled by UCR sequences and are independent of the coding sequence of a gene and its location within the pGKL VLE system.

Synthesis of 5' poly(A) leaders of the poxviral late transcripts is caused by the RNA polymerase slippage on the three consecutive thymidines within the TSS site (Schwer et al., 1987; Schwer and Stunnenberg, 1988; Davison and Moss, 1989b). We showed recently that similar sequence motif is present in TSS of pGKL genes and its alteration can affect 5' polyadenylation of the pGKL mRNAs (Sýkora et al., 2018). We analyzed nucleotide composition directly adjacent to the 5' non-templated adenosines of all pGKL1/2 mRNAs. This analysis revealed that number of 5' non-templated adenosines significantly increased with the number of directly contiguous consecutive adenosine nucleotides coded by the DNA template (Figure 6, Table S8). All these results suggested pGKL-specific RNA polymerase slippage on consecutive thymidine residues of the coding DNA

strand at TSS, which seems to be pronounced by an increased number of coded consecutive adenosines (Figure 6, Table S8). The latter is in contrast to poxviral transcripts, where significant differences in the length of the 5' poly(A) leaders between intermediate and late mRNAs could not be attributed to differences in number of coded consecutive adenosines at the TSS because usually only three consecutive adenosines are coded at their TSS (Yang et al., 2012).

Then, we wanted to investigate our finding that length of the 5' poly(A) leaders of pGKL mRNAs inversely correlates with proportion of their 5' capping (Figure 3) in more detail. We analyzed 5' cap occurrence in transcripts of the *G418<sup>R</sup>* gene which was inserted under the control of the *KIUCR2* within the engineered pRKL1-1 linear VLE and its pRKL1-2 and pRKL1-3 derivatives. The TAAAT sequence of the *KIUCR2* TSS was modified to TAACAT and TACCAT in pRKL1-2 and pRKL1-3, respectively. We showed in our parallel study that mutations in *KIUCR2* TSS led to the shortening, in case of TAACAT TSS, or even loss, in case of TACCAT TSS, of 5' poly(A) leaders of the *G418<sup>R</sup>* transcripts (Sýkora et al., 2018). In accordance with our global analysis of all pGKL transcripts (Figure 3), progressive shortening and loss of 5' poly(A) mRNA leaders caused by incremental substitutions of A to C within *KIUCR2* TSS (Sýkora et al., 2018) led to the progressive increase of 5' cap occurrence from 3.6% of the 5' capped mRNAs starting from the wild-type TSS to 64.3% of 5' capped mRNAs starting from the two-points mutated pRKL1-3 TSS (TACCAT, Figure 7).

We can speculate that pGKL transcription machinery core complex comprising RNA polymerase subunits (K2Orf6p and K2Orf7p) and mRNA capping enzyme (K2Orf3p) (Sýkora et al., 2018) undergo conformational changes influenced by the RNA polymerase slippage at the TSS region that can lead to the decreased frequency of 5' mRNA cap synthesis. Interestingly, nucleotide substitutions and deletions located just downstream of the essential TAAAT motif within the initiator region of the vaccinia virus late promoters led to the variation in lengths of the 5' poly(A) leaders of the corresponding viral mRNAs (de Magistris and Stunnenberg, 1988). This result also suggests that other parts of the pGKL promoters than TSS alone might influence length of the 5' poly(A) leaders and frequency of the 5' mRNA capping. In contrast to our results showing frequent occurrence of the pGKL transcripts lacking 5' caps (Table S4), both vaccinia virus early and postreplicative transcripts are assumed to be fully 5' capped (Kates and Beeson, 1970; Wei and Moss, 1974; 1975; Boone and Moss, 1977).

Our analyses of the pGKL and pRKL2-1 mRNA ends showed the importance of the UCR region for 5' mRNA end formation and may help to explain the surprising results of Schickel *et al.*,

who used a glucose dehydrogenase (*gdh*) reporter gene integrated into the yeast cytoplasmic linear VLEs to measure the strength of several pGKL promoters. They showed that *K2UCR6* controlling the *gdh* reporter resulted in almost two times higher expression than when the reporter was controlled by *K2UCR10* even though these two UCRs have exactly the same length and UCS sequence (Schickel et al., 1996). We found that 20% of *K2ORF6* transcripts possessed a 5' cap, whereas only uncapped mRNAs were found among *K2ORF10* transcripts (Table S4 and Table S5A). A 5' cap positively affects the translation of mRNAs carrying a 5' poly(A) leader in a wheat germ extract (WGE) *in vitro* (Gudkov et al., 2005). If a 5' cap substantially increases the translation efficiency of pGKL transcripts, differences in cap occurrence frequency can account for differences in the expression of pGKL genes.

#### **pGKL mRNAs do not contain 3' poly(A) tails**

The unique 5' ends of the pGKL mRNAs immediately raised a question about the organization of their 3' ends. The pGKL mRNAs can be readily purified by oligo(dT) columns which might be an indication of polyadenylation at their 3' ends or, alternatively, it may be a result of their A/T rich nature (Sor and Fukuhara, 1985; Stark et al., 1990). Possible 3' polyadenylation is reasonable especially in the case of non-capped pGKL mRNAs, because a 3' poly(A) tail of eukaryotic mRNAs plays an important role in their stability (Santiago et al., 1987; Bernstein et al., 1989) and the efficiency of their utilization by cellular translation apparatus (Munroe and Jacobson, 1990; Gallie, 1991; Tanguay and Gallie, 1996). However, pGKL VLEs likely do not encode their own poly(A) polymerase and it remains unknown, whether host cellular polyadenylation machinery participates on processing of pGKL-specific 3' mRNA ends. To answer this question, we performed a modified 3' RACE where a purified and DNase I-treated total yeast RNA was 3' tailed by an *E. coli* poly(A) polymerase using CTP instead of ATP as a substrate. Following reverse transcription with an oligo(dG) anchored primer, PCR amplifications were performed with a universal anchor primer and appropriate gene-specific primers (for primer sequences refer to Table S1). At least twenty randomly selected 3' RACE fragments cloned into pCR4-TOPO plasmids were sequenced for each ORF. The 3' RACE analyses of all 15 pGKL ORFs are exemplified by the results of *K1ORF1* (Figure 2B and Table S5C), *K2ORF5* and *K2ORF10* (Table S5C).

We found that pGKL mRNAs do not contain a 3' poly(A) tail. Transcripts produced from the pGKL VLEs contain heterogeneous 3' ends, suggesting the existence of more than one independent putative termination signal. Transcripts corresponding to only four ORFs are uniform at their 3' ends.

490 This group is exemplified by *K2ORF10* (Table S5C). Four pGKL ORFs gave rise to transcripts that  
 491 fell into two distinct groups with regard to their 3' UTR lengths, suggesting two independent  
 492 transcription terminators. These transcripts are exemplified by those of *K2ORF5* (Table S5C).  
 493 Transcripts of the remaining seven pGKL ORFs probably utilize three or more termination sites.  
 494 Weak RNA stem loops at the 3' ends of the pGKL mRNAs serve as transcription termination signals.  
 495 We presented their detailed analysis recently (Sýkora et al., 2018). Collectively, these results suggest  
 496 that cellular 3' pre-mRNA cleavage and polyadenylation machinery does not act on pGKL mRNAs.  
 497 However, pGKL transcripts can be engineered to contain 3' poly(A) tails and recently, it was shown  
 498 that genetically encoded 3' poly(A) tail enhance the expression of reporter gene located on the pGKL  
 499 VLE (Zhong et al., 2018).

500 **Cellular mRNAs outcompete pGKL mRNAs in binding to yeast cap-binding protein eIF4E *in***  
 501 ***vitro*.**

502 The presence of a methylguanosine cap at their 5' end and 3' polyadenylation are common hallmarks  
 503 of eukaryotic mRNAs. Absence of these structures in the pGKL transcripts should lead to their  
 504 decreased stability and low translatability (Gallie, 1991; Tarun and Sachs, 1995; Tanguay and Gallie,  
 505 1996). However, this is clearly not the case for pGKL VLEs which are stably maintained in yeast  
 506 cells and their genes thus have to be expressed to the level ensuring autonomous VLE replication in  
 507 the host cytoplasm. Previous experiments showed that pGKL mRNAs can be easily detected by  
 508 Northern blots (Stark et al., 1990). Schickel *et al.* investigated strength of the pGKL UCRs  
 509 (promoters) using bacterial glucose dehydrogenase gene (*gdh*) as a reporter. They found that *gdh*  
 510 expression from the yeast linear cytoplasmic VLEs is weaker when compared to bacterial systems  
 511 and yeast nuclear promoters. However, they were able to detect glucose dehydrogenase activity for  
 512 all 8 pGKL UCRs tested (Schickel et al., 1996). It should also be noted, that the toxin encoded by the  
 513 pGKL1 VLE is secreted in high amounts to the culture medium ((Stark and Boyd, 1986) and next  
 514 paragraph) despite the fact that transcripts coding for the toxin subunits and immunity have the  
 515 lowest cap occurrence frequency among all of the pGKL transcripts (*K2ORF2* - 7.5%, *K2ORF3* -  
 516 5.3%, *K2ORF4* - 17.4% of capped mRNAs; Table S4 and Table S5A). To investigate this  
 517 phenomenon, we designed a series of *in vitro* and *in vivo* experiments in which we have  
 518 demonstrated that killer toxin mRNAs, and probably other pGKL mRNAs with uncapped 5' poly(A)  
 519 leaders are translated in an eIF4E-independent manner.



Translation initiation factor 4E (eIF4E) plays a crucial role in eukaryotic translation initiation. It recognizes and binds m<sup>7</sup>G cap structures at the 5' ends of the eukaryotic mRNAs and brings them to the initiating small ribosomal subunits *via* its interaction with scaffold factor eIF4G. The eIF4E protein is essential in all eukaryotes, and its deletion is also lethal in yeast (Altmann et al., 1987; Brenner et al., 1988; Giaever et al., 2002). The structure and function of the cap-binding eIF4E protein from the yeast *S. cerevisiae* is very well understood. Because pGKL VLEs can be transferred and stably maintained in *S. cerevisiae* (Gunge and Sakaguchi, 1981), we decided to test the strength of interaction between the *S. cerevisiae* eIF4E cap-binding protein (S.c.-eIF4E) and pGKL mRNAs in the presence of excess of total RNA *in vitro*.

We produced S.c.-eIF4E as an N-terminal GST-fusion protein in *E. coli* and purified it by glutathione-Sepharose affinity chromatography. To test an interaction of pGKL-specific mRNAs with the yeast eIF4E in the presence of excess of cellular RNA, we incubated DNase I-treated total RNA purified from *K. lactis* IFO1267 with the purified GST-S.c.-eIF4E fusion protein bound to the glutathione Sepharose. After 3 hours of incubation at room temperature, the slurry was extensively washed six times with a great excess of buffer I at room temperature and directly used for RT-PCRs with gene-specific primers. Figure 8 clearly shows that the *K2ORF5* mRNA was completely washed out from the GST-S.c.-eIF4E Sepharose in the first two washing steps (Figure 8, lanes 3 and 4) whereas cellular mRNAs, represented by mRNA coding for a high-affinity glucose transporter (*HGT1*) remained bound to the eIF4E even after six thorough washing steps (Figure 8, lanes 6 and 7). This result clearly indicates that, contrary to *HGT1* mRNA, *K2ORF5* mRNA does not bind to yeast eIF4E *in vitro* and/or binds to it with low affinity and, therefore, can be outcompeted with an excess of yeast cellular mRNAs. This result is further supported by qRT-PCR analysis that revealed comparable abundances of *K2ORF5* and *HGT1* mRNAs in the *K. lactis* IFO1267 total RNA (Figure S3).

Interestingly, *K2ORF5* does not belong to the group of the pGKL mRNAs with the lowest frequency of 5' m<sup>7</sup>G cap occurrence (Table S4 and Table S5A). An explanation could be that the strength of eIF4E binding to mRNA is highly influenced by the first proximal nucleotide following the terminal m<sup>7</sup>G, where the association constant  $K_a$  is much higher for purines than for pyrimidines and slightly higher for guanosine than for adenosine (Wieczorek et al., 1999; Tamarkin-Ben-Harush et al., 2017). Many similar results have been obtained by other researchers; however, cap analogs were used to determine the  $K_a$  and  $K_d$  values in most of these studies. Presence of more nucleotides after the first proximal nucleotide substantially increased eIF4E binding to the capped RNA molecule



while still preserving the differential influence of the first proximal nucleotide on eIF4E binding (Tamarkin-Ben-Harush et al., 2017). A short homopolymeric stretch of 5' adenosine nucleotides thus might destabilize eIF4E binding to the mRNA cap. However, decreased binding of the pGKL mRNAs to eIF4E in competition with an excess of cellular RNA *in vitro* does not substantially affect the performance of those mRNAs in translation *in vivo*, as evidenced by stable maintenance of the pGKL1/2 VLEs in the cell, which suggests an unusual mechanism of translation initiation independent of eIF4E cap-binding protein.

#### **Killer toxin encoded by pGKL VLEs is translated by eIF4E-independent pathway**

The fact that pGKL mRNAs were outcompeted by cellular mRNAs in binding to yeast cap-binding protein eIF4E *in vitro* prompted us to test whether these mRNAs required a functional eIF4E-dependent pathway for their translation initiation. The reliable determination of whether a specific mRNA is translated eIF4E-independently *in vivo* is complicated (Mokrejs et al., 2010). We employed yeast strains bearing conditional mutations in the *CDC33* gene, which codes for eIF4E. *S. cerevisiae* CWO4 strains carrying either a *cdc33-1* or *cdc33-42* allele as a sole source of the eIF4E translation initiation factor (kindly obtained from Prof. Michael Altman) show a strong drop in protein synthesis (Altmann et al., 1989; Altmann and Trachsel, 1989) and a remarkable decrease in the number of polysomes (Feketova et al., 2010) at an elevated temperature. We decided to transfer pGKL VLEs into these strains and determine whether a reasonable amount of the functional pGKL toxin is produced at 37°C, where Cdc33-1 and Cdc33-42 proteins are not functional and eIF4E-dependent translation initiation is thus impaired (Altmann et al., 1989; Altmann and Trachsel, 1989). At least three genes (*KIORF2*, *KIORF3* and *KIORF4*) encoded by the pGKL1 VLE are required to be sufficiently expressed in the yeast cells to impart killer and immunity phenotypes. The killer toxin itself is a trimeric protein comprising non-identical alpha, beta and gamma subunits that are encoded by *KIORF2* and *KIORF4* (Figure 1) (Sugisaki et al., 1984; Stark and Boyd, 1986). Interestingly, transcripts of all these three genes belong to the mRNA group with the lowest cap occurrence frequency and a high degree of 5' polyadenylation (Table S4, Table S5A and Figure 4). The presence of the killer phenotype under restrictive conditions, when cellular eIF4E-dependent translation initiation is impaired, would thus indicate that mRNAs of at least three pGKL genes can be sufficiently and co-ordinately expressed in an eIF4E-independent manner.

The artificial transfer of pGKL VLEs naturally occurring in *K. lactis* into *S. cerevisiae* is rare and difficult. In addition to other obstacles, the host mitochondria and pGKL VLEs in *S. cerevisiae*

are incompatible, and the recipient strain thus must be devoid of mitochondrial DNA ( $\rho^0$ ) (Gunge and Yamane, 1984). We prepared CWO4  $\rho^0$  strains bearing either wild-type *CDC33* or its temperature-sensitive *cdc33-1* or *cdc33-42* alleles by their long-time cultivation in the presence of ethidium bromide. To facilitate the transfer of pGKL VLEs to the new strains and to preserve their isogenic genetic background, we performed cytoduction using an incomplete mating (Zakharov et al., 1969) of the CWO4  $\rho^0$  strains with a nuclear fusion-deficient (*kar1*) yeast strain (*S. cerevisiae* YAT547) carrying both pGKL1 and pGKL2 VLEs (Gunge and Yamane, 1984). We obtained isogenic  $\rho^0$  and [pGKL1/2] *S. cerevisiae* strains on the CWO4 genetic background that differed only in their *CDC33* alleles. All these strains contain both pGKL VLEs and display a killer phenotype (Figure 9A).

We required a gene coding for a yeast single-subunit protein toxin to serve as a control for eIF4E-dependent translation. For this purpose, we chose a thoroughly studied K1 killer toxin encoded by the M1 viral satellite dsRNA from the yeast *S. cerevisiae*. We prepared the M1 cDNA, cloned a gene coding for the prepro-K1 killer toxin into a yeast shuttle expression plasmid (Valis et al., 2006) under the control of the strong constitutive *TPI* promoter (plasmid pYX212::M1) and introduced this expression plasmid into the all three *S. cerevisiae* CWO4  $\rho^0$  strains carrying different *CDC33* alleles as described above. The resulting strains display K1 killer and immunity phenotypes at the permissive temperature (24°C) (Figure 9B).

The experiment depicted in Figure 9 shows a comparison of the production of the *K. lactis* toxin expressed naturally from the pGKL1 VLE and the K1 toxin expressed artificially from the nuclear plasmid pYX212::M1. The K1 toxin gene was transcribed by RNA polymerase II in the nucleus and thus provided a 5' m<sup>7</sup>G-capped and 3' polyadenylated mRNA control. We cultivated all the CWO4  $\rho^0$  [pGKL1/2] and CWO4  $\rho^0$  [pYX212::M1] strains in parallel to the exponential phase at the permissive temperature (24°C). All the strains were then harvested and briefly washed three times with an excess of the fresh medium preheated to 24°C. Half of each culture was further cultivated at 24°C. The second half was rapidly transferred to a fresh medium preheated to 37°C and further cultivated at this elevated temperature. At this point, initial samples marked zero were taken for biomass measurements and the determination of toxin activity in the culture medium. Toxin activities were undetectable in all samples at time zero and provided clear evidence that all cultures were washed sufficiently and that the toxin activities present in samples taken after 3, 6 and 12 hours of cultivation corresponded to proteins produced only in the course of the experiment (Figure 9).

Importantly, the production of the *K. lactis* killer toxin in *S. cerevisiae* strains bearing pGKL VLEs was almost comparable under permissive and restrictive conditions, regardless of whether the

wild-type or temperature-sensitive *CDC33* allele was present (Figure 9A; 37°C). Production of the K1 toxin in the *cdc33-1* and *cdc33-42* yeast strains containing the pYX212::M1 plasmid vanished rapidly at 37°C, whereas the wild-type Cdc33p cap-binding protein permitted K1 toxin production at both 24°C and 37°C (Figure 9B). The presence of the temperature-sensitive eIF4E proteins in *cdc33-1* and *cdc33-42* yeast strains was documented by their restricted growth at 37°C. Both strains bearing the wild-type allele of the *CDC33* gene grew well at both 24°C and 37°C (Figure 9). The growth of CWO4 *CDC33*,  $\rho^0$  [pYX212::M1] was slightly slower at 24°C than at 37°C. The most likely reason is that K1 killer toxin is more active at 24°C than at 37°C, even though we cloned the coding sequence of the K1 super-killer variant that shows enhanced thermal stability (Palfree and Bussey, 1979). We previously showed that K1 preprotoxin produced artificially from the nuclear plasmid does not fully support the immunity phenotype of the host cells (Valis et al., 2006). Higher K1 toxin stability at 24°C thus probably led to the slightly slower growth. However, both cultures eventually reached comparable biomass, in contrast to the strains bearing conditional *cdc33<sup>ts</sup>* alleles cultured at a restrictive temperature (37°C), which completed the initiated cell division and ceased growth (Figure 9B). The activity of the *TPI* promoter used to produce K1 toxin from the pYX212::M1 plasmid declines at the end of exponential growth (Partow et al., 2010), and this phenomenon is even more profound in the version of the *TPI* promoter used in the pYX212 plasmid (Novák, 2012). The growth of the CWO4 *CDC33*,  $\rho^0$  [pYX212::M1] strain was slightly faster at 37°C than at 24°C, causing the culture to reach early stationary phase in 8 hours (instead of 12 at 24°C), and this difference explains why K1 toxin production ceased in the *CDC33* wild-type strain grown for 12 hours at 37°C (Figure 9B).

The most straightforward explanation of the results obtained in this experiment is that whereas K1 toxin mRNA was transcribed from the shuttle plasmid under the control of the *TPI* promoter by the nuclear RNA polymerase II and thus it was 5' capped, 3' polyadenylated, and translated apparently only by the canonical eIF4E-dependent pathway, pGKL1 toxin mRNAs retained their ability to be translated even when 5' mRNA cap recognition by the cellular translation initiation machinery was severely impaired. The pGKL1 mRNAs are transcribed by the pGKL VLE-specific RNA polymerase encoded by *K2ORF6/7* and contain unusual 5' stretches of non-template adenosine nucleotides rarely covalently linked to the terminal m<sup>7</sup>G caps (Table S4, Table S5A, Figures 1, 2A, 3 and 4). Our experiments, both *in vitro* (Figure 8), and *in vivo* (Figure 9), strongly suggest that these 5' terminal structures of the pGKL mRNAs support a novel manner of translation initiation independent of the active eIF4E protein.

To investigate whether 5' polyadenylated mRNAs, which do not bear 5' caps and 3' poly(A) tails are actively translated, we decided to analyze their possible presence with actively translating polysomes. Polysome profile analysis was performed as described previously (Masek et al., 2011; Pospisek and Valasek, 2013). We found that these mRNAs were present in polysomes in both the yeast *K. lactis* IFO1267 naturally bearing pGKL1/2 VLEs and the engineered *S. cerevisiae* CWO4 *cdc33-1*  $\rho^0$  [pGKL1/2] strain (Figure 10). The latter allowed us to study an impact of short 5' poly(A) leaders on the mRNA load to polysomes in the elevated temperature when eIF4E-dependent translation initiation was impaired. Figure 10A and C clearly shows that uncapped and 5' polyadenylated mRNAs are readily present in *K. lactis* polysomes under normal growth conditions. Representation of 5' m<sup>7</sup>G capped *K1ORF2* mRNA coding for  $\alpha/\beta$  toxin subunits in *K. lactis* polysomes mirrors its representation in the total RNA preparation and corresponds to 7.5% and 7.6% of the sequenced *K1ORF2* mRNA UTRs, respectively. In the case of *K2ORF5* mRNA, frequency of capped mRNAs is slightly higher in *K. lactis* polysomes in comparison to total mRNA and represents 57% and 34% of the sequenced *K2ORF5* mRNA 5' UTRs, respectively (Figure 10C and Table S4). The *S. cerevisiae* CWO4 *cdc33-1*  $\rho^0$  [pGKL1/2] strain cultured at non-permissive 37°C for six hours revealed typically severely decreased polysomes (Feketova et al., 2010) that nevertheless contained both capped and uncapped *K1ORF2* and *K2ORF5* mRNAs (Figure 10 B and C). Regardless a low mRNA content, we succeeded to analyze few 5' UTRs of *K1ORF2* and *K2ORF5* mRNAs from the decreased polysomes. The overall ratio between mRNAs containing 5' poly(A) leaders and m<sup>7</sup>G caps retained the same trend both in the polysomes from *S. cerevisiae* cells lacking functional eIF4E, and in the polysomes from *K. lactis* cultivated under normal growth conditions, suggesting of a similar organisation of pGKL-specific 5' mRNA ends in both species. Interestingly, we could not detect any polysomal *K2ORF8* mRNA in the absence of the functional cap-binding protein eIF4E (Figure 10C). This is not in agreement with the presumed cellular level of *K2ORF8* mRNA based on the strength of the pGKL1/2 promoters measured by Schickel *et al.*, who showed that the strength of *K2ORF8* and *K2ORF5* promoters is almost equal, whereas *K1ORF2* promoter is  $\approx 2.5$  times weaker in comparison to both *K2ORF8* and *K2ORF5* promoters (Schickel et al., 1996). Based on these data we speculate that loading of nearly fully capped pGKL mRNAs with a very short or missing 5' poly(A) leader (*K2ORF2*, *K2ORF3* and *K2ORF8*) onto polysomes might be considerably decreased in cells lacking functional eIF4E.

Short uninterrupted 5' poly(A) leaders have been recently reported to increase translation efficiency of the reporter mRNAs in cells infected with the vaccinia virus, including cells exhibiting

decreased eIF4E levels mediated by siRNA knock-down (Dhungel et al., 2017). The average length of the 5' poly(A) leaders of the intermediate and late mRNAs of the vaccinia virus is between 8 and 12 adenosine nucleotides (Yang et al., 2012). The 5' poly(A) leaders of this length also appeared to be most efficient in translation of the reporter mRNAs transfected into HeLa cells infected with the vaccinia virus, whereas 5' leaders comprising only four consecutive adenosines were translationally inefficient (Dhungel et al., 2017).

We analyzed an average length of the uninterrupted 5' poly(A) leaders consisting from both templated and non-templated adenosine nucleotides in all pGKL1/2 transcripts (Figure 4 and Table S4). The average number of consecutive adenosine nucleotides at the very 5' end is 10 (including AUG) and 5.7 for *K1ORF2* and *K2ORF5* mRNAs, respectively. Conversely, frequently capped *K2ORF8* mRNAs contain only a short 5' leader consisting of two templated and one non-templated adenosine nucleotides in average (Figure 4 and Table S4).

Collectively, these results (i) strongly suggest that loading of pGKL1/2 mRNAs into polysomes is mediated by both 5' poly(A) leaders and m<sup>7</sup>G cap depending on the length of the 5' consecutive adenosine nucleotide stretch, (ii) demonstrate that pGKL1/2 mRNAs bearing longer 5' poly(A) leaders can be translated in eIF4E-independent manner, (iii) indicate that pGKL mRNAs bearing very short 5' poly(A) leader of 2-3 nts require functional eIF4E to be utilized by the translational machinery and (iv) explain production of the killer toxin coded by *K1ORF2* and *K1ORF4* in an absence of the functional eIF4E protein in yeast cells.

#### **Pab1 and Lsm1 deletions support translation of pGKL mRNAs**

Poly(A) tracts within the 5' UTRs and the 5' poly(A) leaders of eukaryotic and viral mRNAs have been identified as enhancers of mRNA translation in various examples (Gudkov et al., 2005; Gilbert et al., 2007; Shirokikh and Spirin, 2008; Tahiri-Alaoui et al., 2014) as well as translational repressors in other cases (Melo et al., 2003; Patel et al., 2005). Some studies even show that poly(A) stretches within 5' UTRs may serve as translational enhancers until they reach a certain threshold length. Further increase in the number of consecutive adenosine nucleotides beyond the threshold length leads to translational repression (Xia et al., 2011). Poly(A)-binding protein 1 (PABP1) plays an important role in many of these studies and biological examples. PABP1 is an abundant general translation initiation factor that binds to 3' poly(A) tails, a common mark of most eukaryotic mRNAs (Sachs et al., 1986). In addition, PABP1 serves as a specific regulatory protein and has an important role in mRNA stability and turnover. The importance of PABP1 in cellular translation and other vital



processes is underlined by increasing evidence that it is targeted by a wide range of viruses aiming to take over host protein synthesis or to escape the cellular antiviral defence (Kahvejian et al., 2005; Smith and Gray, 2010; Kobayashi et al., 2012; Costello et al., 2015).

To investigate the possible effects of the yeast PABP1 protein (Pab1) on the expression of pGKL genes, we deleted the *PAB1* gene from *K. lactis*. Because no experimental data describing *PAB1* deletion in the yeast *K. lactis* were available, we followed a strategy used for the deletion of *PAB1* in the yeast *S. cerevisiae* (Mangus et al., 1998). Deletion of *PAB1* gene is lethal in *S. cerevisiae* (Sachs et al., 1987; Giaever et al., 2002) but this lethality can be suppressed by simultaneous deletion of the gene coding for Pab1p-binding protein (Pbp1) (Mangus et al., 1998). We disrupted the *PBP1* gene in the chromosomal DNA of the *K. lactis* IFO1267 strain using a loxP-*G418*-loxP system (Guldener et al., 1996) and obtained the *K. lactis* IFO1267 *pbp1::G418* strain. The *G418* cassette was removed from the genome using an artificially produced Cre-recombinase, providing the strain *K. lactis* IFO1267 *pbp1Δ*. We used the same strategy to introduce a *PAB1* deletion and to obtain the required strain, *K. lactis* IFO1267 *pbp1Δ pab1Δ*.

Another protein that stabilizes some viral 5' poly(A) tracts and thus may also affect translation of the respective mRNAs is Lsm1 (Bergman et al., 2007). We disrupted the *LSM1* gene in the *K. lactis* IFO1267 genome using a strategy similar to that described for *PBP1* deletion. The *K. lactis* *pbp1Δ pab1Δ* double-mutant strain is viable but exhibits a slow-growth phenotype, similar to the behaviour reported for *S. cerevisiae* *pbp1Δ pab1Δ* (Mangus et al., 1998). Deletions of *LSM1* and *PBP1* from *S. cerevisiae* decrease growth rates (Mayes et al., 1999; Yoshikawa et al., 2011). In contrast, we detected only minor decrease in the growth rate of the *K. lactis* *lsm1Δ* strain and no difference in the growth rate of the *K. lactis* *pbp1Δ* strain from the growth rate of their parental wild-type strain *K. lactis* IFO1267 (Figure 11).

To detect a possible effect of the deleted proteins on the expression of pGKL mRNAs, we decided to measure *K. lactis* killer toxin production. Production of the toxin to the culture medium was quantified by a well test on the lawn of the sensitive yeast strain. Using a serial dilution of the toxin-containing culture medium and a calibrated digital microscope for measuring a width of the inhibition zone, we have substantially improved range, accuracy and reproducibility of the killer toxin well test. Relationship between the killer toxin concentration and its activity calculated as a width of the inhibition zone normalized to concentration of production cells (I) yields a hyperbolic function, limit of which corresponds to the calculated theoretical width of the normalized inhibition zone at killer toxin saturation ( $I_s$ ) (Material and Methods and Figure S4).



We calculated  $I_s$  for the toxin activity present in the culture medium of all the strains after 35 hours of cultivation. Disruption of *LSM1* and *PAB1*, but not *PBP1* genes enhanced production of the toxin encoded by the linear cytoplasmic pGKL1 VLE (Figure 11 and Figure S4). The Lsm1 protein is a part of the complex of cellular mRNA decapping activators employed by positive-strand RNA viruses to promote their translation and replication (reviewed in (Jungfleisch et al., 2016)) and might stabilize orthopoxviral mRNAs containing 5' poly(A) leaders (Bergman et al., 2007). Despite many similarities between poxviral and pGKL-encoded transcription and RNA modification apparatuses (Jeske et al., 2007), the simplest explanation for increased pGKL toxin production in the *K. lactis* *lsm1Δ* strain might be the stabilization of pGKL mRNAs in the absence of Lsm1p due to prominent involvement of the Lsm1 protein in a cytoplasmic deadenylation-dependent mRNA decay pathway (reviewed in (Parker, 2012)).

We calculated the length of the continuous 5' poly(A) leaders of the pGKL mRNAs while considering both non-templated and template-coded adenosine nucleotides. Even though the maximal length of the 5' mRNA poly(A) leaders can reach 26 nucleotides, their average and median lengths are much lower. The median lengths of 5' poly(A) leaders in transcripts from only two genes, *K1ORF3* and *K2ORF7*, are 12 adenosine nucleotides. Transcripts from *K1ORF1*, *K1ORF2* and *K2ORF11* have 5' poly(A) leaders with median lengths of 10 adenosine nucleotides. All other pGKL transcripts have shorter 5' poly(A) leaders (Figure 4, Table S4 and Table S5A). These results agree very well with the genome-wide study of Xia *et al.*, who showed that the presence of poly(A) stretches in 5' UTRs of *S. cerevisiae* transcripts positively correlates with enhanced protein synthesis until the length of the 5' poly(A) regions reaches at least 12 consecutive adenosine nucleotides (Xia et al., 2011), which is also the optimal length for the binding site of yeast Pab1 (Sachs et al., 1987). Similarly Dhungel *et al.* demonstrated that reporter mRNAs transfected to the HeLa cells infected with vaccinia virus yielded highest translation rate with 5' leaders consisting of 8, 10 and 12 consecutive adenosine residues, where the 12(A) mRNA 5' leader was the most efficient (Dhungel et al., 2017).

The Pab1 protein might function as a negative regulator that can, due to the documented sharp increase in its binding affinity to poly(A) stretches ranging from 8 to 12 adenosine residues in length (Sachs et al., 1987), further modulate the expression of pGKL genes. Such Pab1 activity was suggested for larger set of yeast genes and can be demonstrated for cyclin *PCL5* coding mRNA, which contains the longest yeast 5' UTR poly(A) region and shows inefficient translation and extremely low ribosomal occupancy (Xia et al., 2011). In mammalian cells, PABP1 binds to the

adenosine-rich elements within the 5' UTR of its own mRNA as a part of the negative autoregulatory complex (Patel et al., 2005). Messenger RNAs containing long uncapped poly(A) leaders (25 nt long) are efficiently translated in WGE lysates and, in contrast to other mRNAs, their translation rate does not decrease at high mRNA concentrations (Gudkov et al., 2005). This behaviour suggests a lower dependence of such mRNAs on translation initiation factors. Indeed, mRNAs containing long 5' poly(A) leaders can efficiently enter translation, even in the absence of the otherwise essential translation factors eIF4F, eIF3 and PABP1 (Shirokikh and Spirin, 2008).

Late poxviral mRNAs containing long 5' poly(A) leaders have a low requirement for eIF4F and perhaps have also a low requirement for PABP1 (Mulder et al., 1998; Xia et al., 2011; Dhungel et al., 2017). Interestingly, pGKL1 mRNAs coding for toxin and immunity phenotypes contain a higher fraction of mRNAs with the 5' poly(A) leaders that are at least 12 residues long (Figure 4, Table S4 and Table S5A). These mRNAs, which perhaps normally have Pab1 bound to their 5' UTRs, may become unblocked and available for translation in the *K. lactis* *pbp1Δ pab1Δ* strain, thus increasing overall pGKL toxin production as we have observed. The control strain, *K. lactis* *pbp1Δ*, showed no detectable growth defects and/or enhancement of pGKL toxin production (Figure 11). To exclude the possibility of changes in the 5' poly(A) leader length and the m<sup>7</sup>G capping frequency of pGKL transcripts in the *pbp1Δ pab1Δ* double-mutant strain, we analyzed 5' ends of *KIORF2* and *KIORF3* mRNAs coding for the toxin  $\alpha/\beta$ -subunit and immunity, respectively, by the 5' RACE. As clearly seen in (Table S9), deletion of the *PAB1* gene does not remarkably affect the structure of the 5' ends of the pGKL mRNAs in *K. lactis* cells.

## Conclusions

We employed pGKL VLEs from the yeast *Kluyveromyces lactis* as a model to investigate and present herein a complex view of the largely unexplored transcriptome of the yeast cytoplasmic linear VLEs. Although previous experiments reported the binding of pGKL mRNAs to an oligo(dT) column (Stark et al., 1990), we showed that the pGKL transcripts are not polyadenylated at their 3' ends; however, they frequently contain uncapped 5' poly(A) leaders that are not complementary to the VLEs genomic DNA and are probably synthesized by the RNA polymerase slippage. The degree of 5' capping and/or 5' polyadenylation is specific to each gene and is controlled by the short UCRs sequences preceding the corresponding AUG start codon. We also showed that the lengths of the 5' poly(A) leaders of the pGKL mRNAs inversely correlates with the proportion of the 5' mRNA capping. The pGKL transcriptome analysis allowed us to refine the description of the pGKL

promoters and revealed new alternative promoters and corresponding transcription start sites for the *K1ORF4*, *K2ORF3* and *K2ORF4* genes.

Translation of pGKL transcripts is independent of eIF4E and Pab1 translation factors. With regard to our results and published data on yeast (Xia et al., 2011) and mammalian (Melo et al., 2003; Patel et al., 2005) mRNAs containing long poly(A) stretches in their 5' UTRs, we suggest that Pab1 and Lsm1 serve as negative regulators of the pGKL mRNAs containing 5' poly(A) leaders. It is tempting to speculate that differences in the median lengths of the 5' mRNA poly(A) leaders of each gene (Table S4 and Table S5A, Figure 4) might reflect the degree of inhibition of the corresponding gene expression by Pab1, which might also be an evolutionary force driving the selection of the appropriate UCRs. Our data are in a good agreement with the average length of 8-12 adenosines in the 5' poly(A) leaders of the poxviral post-replicative mRNAs (Yang et al., 2012). Such length of the 5' poly(A) leaders also maximally enhanced translation of the reporter mRNAs transfected into the HeLa cells infected with the vaccinia virus (Dhungel et al., 2017).

We hypothesize that the small and extremely compact genomes of pGKL VLEs use different levels of 5' capping and 5' polyadenylation of their transcripts to finely tune the expression of each gene to proper levels to facilitate the formation of functional protein complexes and to keep pGKL gene expression in harmony with the host cell. Initial evidence supporting this hypothesis came from comparison of our data with that of Schickel *et al.*, as discussed earlier (Schickel et al., 1996), and from our analyses of load of the pGKL mRNAs into polysomes in normal growth conditions and in cells lacking functional cap-binding translation initiation factor 4E. Our data further support a hypothesis about evolutionary relationship between the family of the cytoplasmic yeast linear VLEs and poxviruses (Jeske et al., 2007; Sýkora et al., 2018). This connection is fascinating because members of the *Poxviridae* family have been found only in vertebrates and arthropods, and some of them cause highly contagious and serious diseases of humans and animals. Parallel studies of yeast cytoplasmic linear VLEs and poxviruses may lead us to the origin of eukaryotic DNA viruses. The pGKL VLEs could thus help us to better understand the basic processes and host cell-virus interactions that cannot be easily studied in mammalian systems.

## AUTHOR CONTRIBUTIONS

VV and MP devised the project and the main conceptual ideas. VV, MS, MP and TM designed and performed the experiments. VV and MS created figures. MP received funding. All authors analyzed the data, discussed the results and contributed to the final manuscript.

## FUNDING

This work was supported by the Czech Science Foundation [grant No. GBP305/12/G034]; the Charles University institutional project [SVV-260426]; and the ELIXIR CZ research infrastructure project MEYS [Grant No: LM2015047] including access to computing and storage facilities.

## ACKNOWLEDGEMENTS

We thank Petra Studnickova, Vlasta Pelechova and Natalie Suchankova (Charles University, Czech Republic) for their help. We thank Prof. Michael Altman (University of Bern, Switzerland), Prof. Norio Gunge (Sojo University), Prof. Mike Kiledjian (Rutgers University), Prof. Liang Tong (Columbia University), and Prof. Beate Schwer (Weill Cornell Medical College) for yeast strains and plasmids.

## Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## REFERENCES

- Ahn, B.Y., and Moss, B. (1989). Capped poly(A) leaders of variable lengths at the 5' ends of vaccinia virus late mRNAs. *J Virol* 63(1), 226-232.
- Altmann, M., Handschin, C., and Trachsel, H. (1987). mRNA cap-binding protein: cloning of the gene encoding protein synthesis initiation factor eIF-4E from *Saccharomyces cerevisiae*. *Mol Cell Biol* 7(3), 998-1003.
- Altmann, M., Sonenberg, N., and Trachsel, H. (1989). Translation in *Saccharomyces cerevisiae*: Initiation Factor 4E-Dependent Cell-Free System. *Molecular and Cellular Biology* 9(10), 4467-4472.
- Altmann, M., and Trachsel, H. (1989). Altered mRNA cap recognition activity of initiation factor 4E in the yeast cell cycle division mutant *cdc33*. *Nucleic Acids Res* 17(15), 5923-5931.
- Bergman, N., Moraes, K.C., Anderson, J.R., Zaric, B., Kambach, C., Schneider, R.J., et al. (2007). Lsm proteins bind and stabilize RNAs containing 5' poly(A) tracts. *Nat Struct Mol Biol* 14(9), 824-831. doi: 10.1038/nsmb1287.
- Bernstein, P., Peltz, S.W., and Ross, J. (1989). The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro. *Mol Cell Biol* 9(2), 659-670.
- Boone, R.F., and Moss, B. (1977). Methylated 5'-terminal sequences of vaccinia virus mRNA species made in vivo at early and late times after infection. *Virology* 79(1), 67-80.
- Brenner, C., Nakayama, N., Goebel, M., Tanaka, K., Toh-e, A., and Matsumoto, K. (1988). CDC33 encodes mRNA cap-binding protein eIF-4E of *Saccharomyces cerevisiae*. *Mol Cell Biol* 8(8), 3556-3559.

- Butler, A.R., White, J.H., and Stark, M.J. (1991). Analysis of the response of *Saccharomyces cerevisiae* cells to *Kluyveromyces lactis* toxin. *J Gen Microbiol* 137(7), 1749-1757.
- Costello, J., Castelli, L.M., Rowe, W., Kershaw, C.J., Talavera, D., Mohammad-Qureshi, S.S., et al. (2015). Global mRNA selection mechanisms for translation initiation. *Genome Biol* 16, 10. doi: 10.1186/s13059-014-0559-z.
- Davison, A.J., and Moss, B. (1989a). Structure of vaccinia virus early promoters. *J Mol Biol* 210(4), 749-769. doi: 10.1016/0022-2836(89)90107-1.
- Davison, A.J., and Moss, B. (1989b). Structure of Vaccinia Virus Late Promoters. *J Mol Biol* 210(4), 771-784. doi: 10.1016/0022-2836(89)90108-3.
- de Magistris, L., and Stunnenberg, H.G. (1988). Cis-acting sequences affecting the length of the poly(A) head of vaccinia virus late transcripts. *Nucleic Acids Res* 16(8), 3141-3156.
- Dhungel, P., Cao, S., and Yang, Z. (2017). The 5'-poly(A) leader of poxvirus mRNA confers a translational advantage that can be achieved in cells with impaired cap-dependent translation. *PLoS Pathog* 13(8), e1006602. doi: 10.1371/journal.ppat.1006602.
- Feketova, Z., Masek, T., Vopalensky, V., and Pospisek, M. (2010). Ambiguous decoding of the CUG codon alters the functionality of the *Candida albicans* translation initiation factor 4E. *FEMS Yeast Res* 10(5), 558-569. doi: 10.1111/j.1567-1364.2010.00629.x.
- Fukuhara, H. (1987). The RF1 gene of the killer DNA of yeast may encode a DNA polymerase. *Nucleic Acids Res* 15(23), 10046.
- Fukuhara, H. (1995). Linear DNA plasmids of yeasts. *FEMS Microbiol Lett* 131(1), 1-9.
- Gallie, D.R. (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev* 5(11), 2108-2116.
- Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., et al. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418(6896), 387-391. doi: 10.1038/nature00935.
- Gietz, R.D., and Woods, R.A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 350, 87-96.
- Gilbert, W.V., Zhou, K., Butler, T.K., and Doudna, J.A. (2007). Cap-independent translation is required for starvation-induced differentiation in yeast. *Science* 317(5842), 1224-1227. doi: 10.1126/science.1144467.
- Gowda, M., Li, H., Alessi, J., Chen, F., Pratt, R., and Wang, G.L. (2006). Robust analysis of 5'-transcript ends (5'-RATE): a novel technique for transcriptome analysis and genome annotation. *Nucleic Acids Res* 34(19), e126. doi: 10.1093/nar/gkl522.
- Gudkov, A.T., Ozerova, M.V., Shiryayev, V.M., and Spirin, A.S. (2005). 5'-poly(A) sequence as an effective leader for translation in eukaryotic cell-free systems. *Biotechnol Bioeng* 91(4), 468-473. doi: 10.1002/bit.20525.
- Gueldener, U., Heinisch, J., Koehler, G.J., Voss, D., and Hegemann, J.H. (2002). A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res* 30(6), e23.
- Guldener, U., Heck, S., Fielder, T., Beinhauer, J., and Hegemann, J.H. (1996). A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res* 24(13), 2519-2524.



- 911 Gunge, N. (1983). Yeast DNA plasmids. *Annu Rev Microbiol* 37, 253-276.
- 912 Gunge, N., and Sakaguchi, K. (1981). Intergeneric transfer of deoxyribonucleic acid killer plasmids,  
913 pGK11 and pGK12, from *Kluyveromyces lactis* into *Saccharomyces cerevisiae* by cell fusion.  
914 *J Bacteriol* 147(1), 155-160.
- 915 Gunge, N., Tamaru, A., Ozawa, F., and Sakaguchi, K. (1981). Isolation and Characterization of  
916 Linear Deoxyribonucleic Acid Plasmids from *Kluyveromyces lactis* and the Plasmid-  
917 Associated Killer Character. *J Bacteriol* 145(1), 382-390.
- 918 Gunge, N., and Tokunaga, M. (2004). Linear DNA plasmids and killer system of *Kluyveromyces*  
919 *lactis*. In: Kück U (ed) *The Mycota II, Genetics and biotechnology (2nd ed)*, Springer-Verlag  
920 Berlin, 199-217.
- 921 Gunge, N., and Yamane, C. (1984). Incompatibility of linear DNA killer plasmids pGKL1 and  
922 pGKL2 from *Kluyveromyces lactis* with mitochondrial DNA from *Saccharomyces cerevisiae*.  
923 *J Bacteriol* 159(2), 533-539.
- 924 Hishinuma, F., Nakamura, K., Hirai, K., Nishizawa, R., Gunge, N., and Maeda, T. (1984). Cloning  
925 and nucleotide sequences of the linear DNA killer plasmids from yeast. *Nucleic Acids Res*  
926 12(19), 7581-7597.
- 927 Huang, B., Johansson, M.J., and Bystrom, A.S. (2005). An early step in wobble uridine tRNA  
928 modification requires the Elongator complex. *Rna* 11(4), 424-436.
- 929 Jablonowski, D., Zink, S., Mehlgarten, C., Daum, G., and Schaffrath, R. (2006). tRNA<sup>Glu</sup> wobble  
930 uridine methylation by Trm9 identifies Elongator's key role for zymocin-induced cell death in  
931 yeast. *Mol Microbiol* 59(2), 677-688.
- 932 Jeske, S., Meinhardt, F., and Klassen, R. (2007). "Extranuclear Inheritance: Virus-Like DNA-  
933 Elements in Yeast," in *Progress in Botany*, eds. K. Esser, U. Löttge, W. Beyschlag & J.  
934 Murata. Springer, Berlin, Heidelberg), 98-129.
- 935 Jeske, S., Tiggemann, M., and Meinhardt, F. (2006). Yeast autonomous linear plasmid pGKL2:  
936 ORF9 is an actively transcribed essential gene with multiple transcription start points. *FEMS*  
937 *Microbiol Lett* 255(2), 321-327.
- 938 Jung, G.H., Leavitt, M.C., and Ito, J. (1987). Yeast killer plasmid pGKL1 encodes a DNA  
939 polymerase belonging to the family B DNA polymerases. *Nucleic Acids Res* 15(21), 9088.
- 940 Jungfleisch, J., Blasco-Moreno, B., and Diez, J. (2016). Use of Cellular Decapping Activators by  
941 Positive-Strand RNA Viruses. *Viruses* 8(12), 340. doi: 10.3390/v8120340.
- 942 Kahvejian, A., Svitkin, Y.V., Sukarieh, R., M'Boutchou, M.N., and Sonenberg, N. (2005).  
943 Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts  
944 via multiple mechanisms. *Genes Dev* 19(1), 104-113. doi: 10.1101/gad.1262905.
- 945 Kamper, J., Meinhardt, F., Gunge, N., and Esser, K. (1989). New recombinant linear DNA-elements  
946 derived from *Kluyveromyces lactis* killer plasmids. *Nucleic Acids Res* 17(4), 1781. doi:  
947 10.1093/nar/17.4.1781.
- 948 Kast, A., Klassen, R., and Meinhardt, F. (2014). rRNA fragmentation induced by a yeast killer toxin.  
949 *Mol Microbiol* 91(3), 606-617. doi: 10.1111/mmi.12481.
- 950 Kates, J., and Beeson, J. (1970). Ribonucleic acid synthesis in vaccinia virus. I. The mechanism of  
951 synthesis and release of RNA in vaccinia cores. *J Mol Biol* 50(1), 1-18. doi: 10.1016/0022-  
952 2836(70)90100-2.



- 953 Klassen, R., and Meinhardt, F. (2007). "Linear Protein-Primed Replicating Plasmids in Eukaryotic  
954 Microbes," in *Microbial Linear Plasmids. Microbiology Monographs*, eds. F. Meinhardt & R.  
955 Klassen. 12 May 2007 ed: Springer, Berlin, Heidelberg), 187-226.
- 956 Kobayashi, M., Arias, C., Garabedian, A., Palmenberg, A.C., and Mohr, I. (2012). Site-specific  
957 cleavage of the host poly(A) binding protein by the encephalomyocarditis virus 3C proteinase  
958 stimulates viral replication. *J Virol* 86(19), 10686-10694. doi: 10.1128/JVI.00896-12.
- 959 Larsen, M., Gunge, N., and Meinhardt, F. (1998). Kluyveromyces lactis killer plasmid pGKL2:  
960 evidence for a viral-like capping enzyme encoded by ORF3. *Plasmid* 40(3), 243-246.
- 961 Lin, R.J., Kim, D.H., Castanotto, D., Westaway, S., and Rossi, J.J. (1996). "RNA preparation from  
962 yeast cells " in *A Laboratory Guide to RNA: Isolation, Analysis, and Synthesis*, ed. P.A.  
963 Krieg. Wiley-Liss, New York), 43-50.
- 964 Lu, J., Huang, B., Esberg, A., Johansson, M.J., and Bystrom, A.S. (2005). The Kluyveromyces lactis  
965 gamma-toxin targets tRNA anticodons. *Rna* 11(11), 1648-1654.
- 966 Mangus, D.A., Amrani, N., and Jacobson, A. (1998). Pbp1p, a factor interacting with Saccharomyces  
967 cerevisiae poly(A)-binding protein, regulates polyadenylation. *Mol Cell Biol* 18(12), 7383-  
968 7396.
- 969 Maruyama, K., and Sugano, S. (1994). Oligo-capping: a simple method to replace the cap structure  
970 of eukaryotic mRNAs with oligoribonucleotides. *Gene* 138(1-2), 171-174.
- 971 Masek, T., Valasek, L., and Pospisek, M. (2011). Polysome analysis and RNA purification from  
972 sucrose gradients. *Methods Mol Biol* 703, 293-309. doi: 10.1007/978-1-59745-248-9\_20.
- 973 Masek, T., Vopalensky, V., Suchomelova, P., and Pospisek, M. (2005). Denaturing RNA  
974 electrophoresis in TAE agarose gels. *Anal Biochem* 336(1), 46-50.
- 975 Mayes, A.E., Verdone, L., Legrain, P., and Beggs, J.D. (1999). Characterization of Sm-like proteins  
976 in yeast and their association with U6 snRNA. *Embo J* 18(15), 4321-4331. doi:  
977 10.1093/emboj/18.15.4321.
- 978 McNeel, D.G., and Tamanoi, F. (1991). Terminal region recognition factor 1, a DNA-binding protein  
979 recognizing the inverted terminal repeats of the pGKI linear DNA plasmids. *Proc Natl Acad*  
980 *Sci U S A* 88(24), 11398-11402.
- 981 Meineke, B., Kast, A., Schwer, B., Meinhardt, F., Shuman, S., and Klassen, R. (2012). A fungal  
982 anticodon nuclease ribotoxin exploits a secondary cleavage site to evade tRNA repair. *Rna*  
983 18(9), 1716-1724. doi: 10.1261/rna.034132.112.
- 984 Meinhardt, F., Wodara, C., Larsen, M., and Schickel, J. (1994). A Novel-Approach to Express a  
985 Heterologous Gene on Kluyveromyces-Lactis Linear Killer Plasmids - Expression of the  
986 Bacterial Aph Gene from a Cytoplasmic Promoter Fragment without in-Phase Fusion to the  
987 Plasmid Open Reading Frame. *Plasmid* 32(3), 318-327. doi: 10.1006/plas.1994.1071.
- 988 Melo, E.O., de Melo Neto, O.P., and Martins de Sa, C. (2003). Adenosine-rich elements present in  
989 the 5'-untranslated region of PABP mRNA can selectively reduce the abundance and  
990 translation of CAT mRNAs in vivo. *FEBS Lett* 546(2-3), 329-334.
- 991 Mokrejs, M., Masek, T., Vopalensky, V., Hlubucek, P., Delbos, P., and Pospisek, M. (2010).  
992 IRESite—a tool for the examination of viral and cellular internal ribosome entry sites.  
993 *Nucleic Acids Res* 38(Database issue), D131-D136. doi: 10.1093/nar/gkp981.

- 994 Mukherjee, C., Patil, D.P., Kennedy, B.A., Bakthavachalu, B., Bundschuh, R., and Schoenberg, D.R.  
995 (2012). Identification of cytoplasmic capping targets reveals a role for cap homeostasis in  
996 translation and mRNA stability. *Cell Rep* 2(3), 674-684. doi: 10.1016/j.celrep.2012.07.011.
- 997 Mulder, J., Robertson, M.E., Seamons, R.A., and Belsham, G.J. (1998). Vaccinia virus protein  
998 synthesis has a low requirement for the intact translation initiation factor eIF4F, the cap-  
999 binding complex, within infected cells. *J Virol* 72(11), 8813-8819.
- 1000 Munroe, D., and Jacobson, A. (1990). mRNA poly(A) tail, a 3' enhancer of translational initiation.  
1001 *Mol Cell Biol* 10(7), 3441-3455.
- 1002 Novák, J. (2012). *Dissecting the nuclear function of interleukin-1alpha*. MSc. Diploma thesis,  
1003 Charles University.
- 1004 Palfree, R.G., and Bussey, H. (1979). Yeast killer toxin: purification and characterisation of the  
1005 protein toxin from *Saccharomyces cerevisiae*. *European Journal of Biochemistry* 93(3), 487-  
1006 493.
- 1007 Parker, R. (2012). RNA degradation in *Saccharomyces cerevisiae*. *Genetics* 191(3), 671-702. doi:  
1008 10.1534/genetics.111.137265.
- 1009 Partow, S., Siewers, V., Bjorn, S., Nielsen, J., and Maury, J. (2010). Characterization of different  
1010 promoters for designing a new expression vector in *Saccharomyces cerevisiae*. *Yeast* 27(11),  
1011 955-964. doi: 10.1002/yea.1806.
- 1012 Patel, D.D., and Pickup, D.J. (1987). Messenger RNAs of a strongly-expressed late gene of cowpox  
1013 virus contain 5'-terminal poly(A) sequences. *Embo J* 6(12), 3787-3794.
- 1014 Patel, G.P., Ma, S., and Bag, J. (2005). The autoregulatory translational control element of poly(A)-  
1015 binding protein mRNA forms a heteromeric ribonucleoprotein complex. *Nucleic Acids Res*  
1016 33(22), 7074-7089. doi: 10.1093/nar/gki1014.
- 1017 Pospisek, M., and Palkova, Z. (1991). Microisolation of Yeast Nucleic-Acids on the Microtitre Plate  
1018 without Using Lytic Enzymes. *Nucleic Acids Research* 19(18), 5083-5083. doi:  
1019 10.1093/nar/19.18.5083.
- 1020 Pospisek, M., and Valasek, L. (2013). Polysome profile analysis--yeast. *Methods Enzymol* 530, 173-  
1021 181. doi: 10.1016/B978-0-12-420037-1.00009-9.
- 1022 Ramanathan, A., Robb, G.B., and Chan, S.H. (2016). mRNA capping: biological functions and  
1023 applications. *Nucleic Acids Res* 44(16), 7511-7526. doi: 10.1093/nar/gkw551.
- 1024 Romanos, M.A., and Boyd, A. (1988). A transcriptional barrier to expression of cloned toxin genes  
1025 of the linear plasmid k1 of *Kluyveromyces lactis*: evidence that native k1 has novel  
1026 promoters. *Nucleic Acids Res* 16(15), 7333-7350.
- 1027 Ross, J. (1995). mRNA stability in mammalian cells. *Microbiol Rev* 59(3), 423-450.
- 1028 Sachs, A.B., Bond, M.W., and Kornberg, R.D. (1986). A single gene from yeast for both nuclear and  
1029 cytoplasmic polyadenylate-binding proteins: domain structure and expression. *Cell* 45(6),  
1030 827-835.
- 1031 Sachs, A.B., Davis, R.W., and Kornberg, R.D. (1987). A single domain of yeast poly(A)-binding  
1032 protein is necessary and sufficient for RNA binding and cell viability. *Mol Cell Biol* 7(9),  
1033 3268-3276.

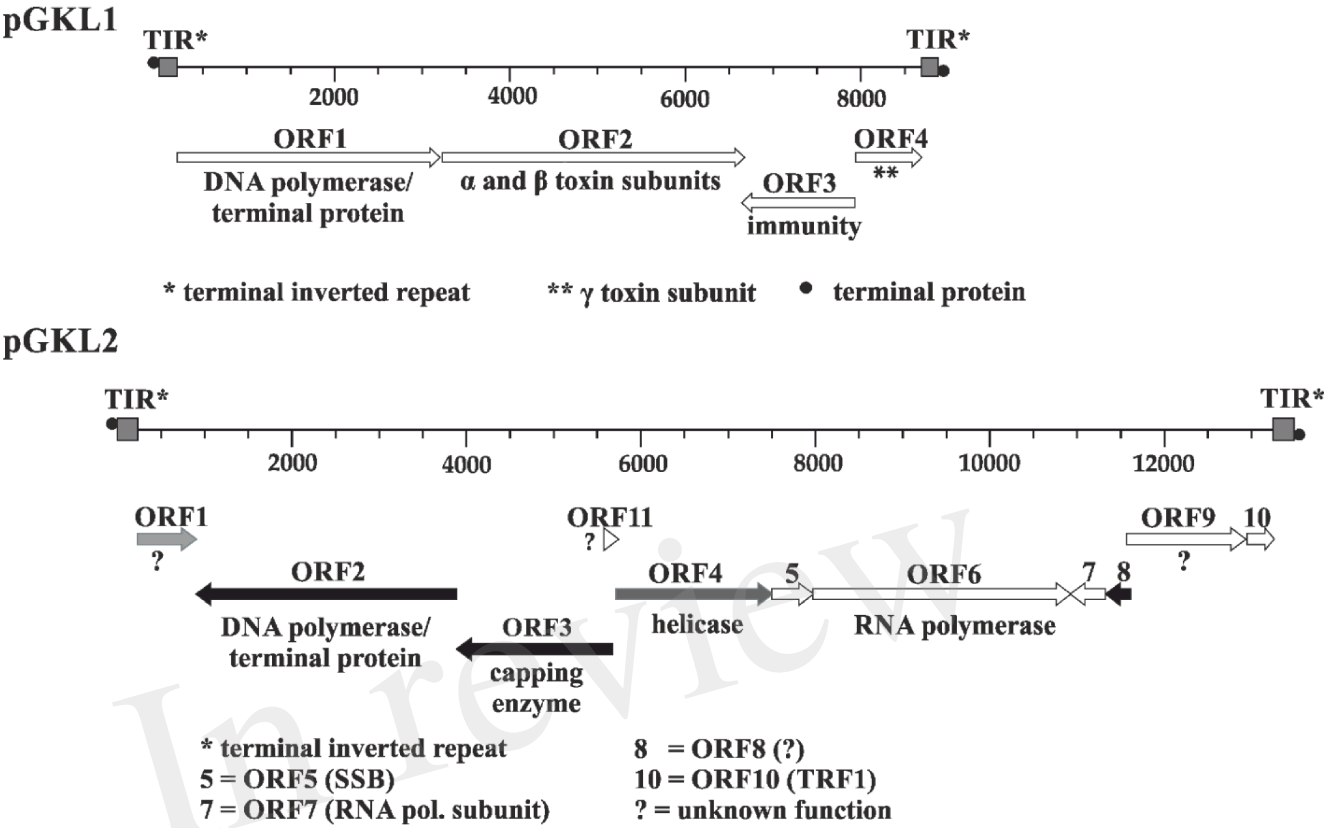
- 1034 Santiago, T.C., Bettany, A.J., Purvis, I.J., and Brown, A.J. (1987). Messenger RNA stability in  
1035 *Saccharomyces cerevisiae*: the influence of translation and poly(A) tail length. *Nucleic Acids*  
1036 *Res* 15(6), 2417-2429.
- 1037 Shirokikh, N.E., and Spirin, A.S. (2008). Poly(A) leader of eukaryotic mRNA bypasses the  
1038 dependence of translation on initiation factors. *Proc Natl Acad Sci U S A* 105(31), 10738-  
1039 10743. doi: 10.1073/pnas.0804940105.
- 1040 Schaffrath, R., and Meacock, P.A. (1995). *Kluyveromyces lactis* killer plasmid pGKL2: molecular  
1041 analysis of an essential gene, ORF5. *Yeast* 11(7), 615-628.
- 1042 Schaffrath, R., and Meacock, P.A. (2001). An SSB encoded by and operating on linear killer  
1043 plasmids from *Kluyveromyces lactis*. *Yeast* 18(13), 1239-1247.
- 1044 Schaffrath, R., Meinhardt, F., and Meacock, P.A. (1997). ORF7 of yeast plasmid pGKL2: analysis of  
1045 gene expression in vivo. *Curr Genet* 31(2), 190-192.
- 1046 Schaffrath, R., Meinhardt, F., and Meacock, P.A. (1999). Genetic manipulation of *Kluyveromyces*  
1047 *lactis* linear DNA plasmids: gene targeting and plasmid shuffles. *FEMS Microbiol Lett*  
1048 178(2), 201-210.
- 1049 Schickel, J., Helmig, C., and Meinhardt, F. (1996). *Kluyveromyces lactis* killer system: analysis of  
1050 cytoplasmic promoters of the linear plasmids. *Nucleic Acids Res* 24(10), 1879-1886.
- 1051 Schmidt, W.M., and Mueller, M.W. (1999). CapSelect: a highly sensitive method for 5' CAP-  
1052 dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nucleic*  
1053 *Acids Res* 27(21), e31.
- 1054 Schrunder, J., Gunge, N., and Meinhardt, F. (1996). Extranuclear expression of the bacterial xylose  
1055 isomerase (xylA) and the UDP-glucose dehydrogenase (hasB) genes in yeast with  
1056 *Kluyveromyces lactis* linear killer plasmids as vectors. *Curr Microbiol* 33(5), 323-330.
- 1057 Schrunder, J., and Meinhardt, F. (1995). An extranuclear expression system for analysis of  
1058 cytoplasmic promoters of yeast linear killer plasmids. *Plasmid* 33(2), 139-151.
- 1059 Schwer, B., Mao, X., and Shuman, S. (1998). Accelerated mRNA decay in conditional mutants of  
1060 yeast mRNA capping enzyme. *Nucleic Acids Res* 26(9), 2050-2057. doi:  
1061 10.1093/nar/26.9.2050.
- 1062 Schwer, B., and Stunnenberg, H.G. (1988). Vaccinia virus late transcripts generated in vitro have a  
1063 poly(A) head. *Embo J* 7(4), 1183-1190.
- 1064 Schwer, B., Visca, P., Vos, J.C., and Stunnenberg, H.G. (1987). Discontinuous transcription or RNA  
1065 processing of vaccinia virus late messengers results in a 5' poly(A) leader. *Cell* 50(2), 163-  
1066 169.
- 1067 Smith, R.W., and Gray, N.K. (2010). Poly(A)-binding protein (PABP): a common viral target.  
1068 *Biochem J* 426(1), 1-12. doi: 10.1042/BJ20091571.
- 1069 Sor, F., and Fukuhara, H. (1985). Structure of a linear plasmid of the yeast *Kluyveromyces lactis*;  
1070 Compact organization of the killer genome. *Current Genetics* 9(2), 147-155. doi:  
1071 10.1007/BF00436963.
- 1072 Stam, J.C., Kwakman, J., Meijer, M., and Stuitje, A.R. (1986). Efficient isolation of the linear DNA  
1073 killer plasmid of *Kluyveromyces lactis*: evidence for location and expression in the cytoplasm  
1074 and characterization of their terminally bound proteins. *Nucleic Acids Res* 14(17), 6871-6884.

- 1075 Stark, M.J., and Boyd, A. (1986). The killer toxin of *Kluyveromyces lactis*: characterization of the  
1076 toxin subunits and identification of the genes which encode them. *Embo J* 5(8), 1995-2002.
- 1077 Stark, M.J., Mileham, A.J., Romanos, M.A., and Boyd, A. (1984). Nucleotide sequence and  
1078 transcription analysis of a linear DNA plasmid associated with the killer character of the yeast  
1079 *Kluyveromyces lactis*. *Nucleic Acids Res* 12(15), 6011-6030.
- 1080 Stark, M.J.R., Boyd, A., Mileham, A.J., and Romanos, M.A. (1990). The Plasmid-encoded Killer  
1081 System of *Kluyveromyces lactis*: A Review. *Yeast* 6(1), 1-29. doi: 10.1002/yea.320060102.
- 1082 Sugisaki, Y., Gunge, N., Sakaguchi, K., Yamasaki, M., and Tamura, G. (1984). Characterization of a  
1083 novel killer toxin encoded by a double-stranded linear DNA plasmid of *Kluyveromyces*  
1084 *lactis*. *European Journal of Biochemistry* 141(2), 241-245. doi: 10.1111/j.1432-  
1085 1033.1984.tb08183.x.
- 1086 Sýkora, M., Pospíšek, M., Novák, J., Mrvová, S., Krásný, L., and Vopálenský, V. (2018).  
1087 Transcription apparatus of the yeast virus-like elements: Architecture, function, and  
1088 evolutionary origin. *PLoS Pathog* 14(10), e1007377. doi: 10.1371/journal.ppat.1007377.
- 1089 Tahiri-Alaoui, A., Zhao, Y., Sadigh, Y., Popplestone, J., Kgosana, L., Smith, L.P., et al. (2014).  
1090 Poly(A) binding protein 1 enhances cap-independent translation initiation of neurovirulence  
1091 factor from avian herpesvirus. *PLoS One* 9(12), e114466. doi: 10.1371/journal.pone.0114466.
- 1092 Takeda, M., Hiraishi, H., Takesako, T., Tanase, S., and Gunge, N. (1996). The terminal protein of the  
1093 linear DNA plasmid pGKL2 shares an N-terminal domain of the plasmid-encoded DNA  
1094 polymerase. *Yeast* 12(3), 241-246.
- 1095 Tamarkin-Ben-Harush, A., Vasseur, J.J., Debart, F., Ulitsky, I., and Dikstein, R. (2017). Cap-  
1096 proximal nucleotides via differential eIF4E binding and alternative promoter usage mediate  
1097 translational response to energy stress. *Elife* 6, e21907. doi: 10.7554/eLife.21907.
- 1098 Tanguay, R.L., and Gallie, D.R. (1996). Translational efficiency is regulated by the length of the 3'  
1099 untranslated region. *Mol Cell Biol* 16(1), 146-156.
- 1100 Tarun, S.Z., Jr., and Sachs, A.B. (1995). A common function for mRNA 5' and 3' ends in translation  
1101 initiation in yeast. *Genes Dev* 9(23), 2997-3007.
- 1102 Tiggemann, M., Jeske, S., Larsen, M., and Meinhardt, F. (2001). *Kluyveromyces lactis* cytoplasmic  
1103 plasmid pGKL2: heterologous expression of Orf3p and proof of guanylyltransferase and  
1104 mRNA-triphosphatase activities. *Yeast* 18(9), 815-825.
- 1105 Tokunaga, M., Kawamura, A., and Hishinuma, F. (1989). Expression of pGKL killer 28K subunit in  
1106 *Saccharomyces cerevisiae*: identification of 28K subunit as a killer protein. *Nucleic Acids Res*  
1107 17(9), 3435-3446.
- 1108 Tokunaga, M., Wada, N., and Hishinuma, F. (1987). Expression and identification of immunity  
1109 determinants on linear DNA killer plasmids pGKL1 and pGKL2 in *Kluyveromyces lactis*.  
1110 *Nucleic Acids Res* 15(3), 1031-1046.
- 1111 Tommasino, M. (1991). Killer System of *Kluyveromyces lactis*: the Open Reading Frame 10 of the  
1112 pGK12 Plasmid Encodes a Putative DNA Binding Protein. *Yeast* 7(3), 245-252. doi:  
1113 10.1002/yea.320070306.
- 1114 Tommasino, M., Ricci, S., and Galeotti, C.L. (1988). Genome organization of the killer plasmid  
1115 pGK12 from *Kluyveromyces lactis*. *Nucleic Acids Res* 16(13), 5863-5878.

- 1116 Valis, K., Masek, T., Novotna, D., Pospisek, M., and Janderova, B. (2006). Immunity to killer toxin  
1117 K1 is connected with the Golgi-to-vacuole protein degradation pathway. *Folia Microbiol*  
1118 (*Praha*) 51(3), 196-202.
- 1119 Wei, C.M., and Moss, B. (1974). Methylation of newly synthesized viral messenger RNA by an  
1120 enzyme in vaccinia virus. *Proc Natl Acad Sci U S A* 71(8), 3014-3018.
- 1121 Wei, C.M., and Moss, B. (1975). Methylated nucleotides block 5'-terminus of vaccinia virus  
1122 messenger RNA. *Proc Natl Acad Sci U S A* 72(1), 318-322.
- 1123 Wieczorek, Z., Niedzwiecka-Kornas, A., Chlebicka, L., Jankowska, M., Kiraga, K., Stepinski, J., et  
1124 al. (1999). Fluorescence studies on association of human translation initiation factor eIF4E  
1125 with mRNA cap-analogues. *Z Naturforsch C* 54(3-4), 278-284.
- 1126 Wilson, D.W., and Meacock, P.A. (1988). Extranuclear gene expression in yeast: evidence for a  
1127 plasmid-encoded RNA polymerase of unique structure. *Nucleic Acids Res* 16(16), 8097-8112.
- 1128 Xia, X., MacKay, V., Yao, X., Wu, J., Miura, F., Ito, T., et al. (2011). Translation initiation: a  
1129 regulatory role for poly(A) tracts in front of the AUG codon in *Saccharomyces cerevisiae*.  
1130 *Genetics* 189(2), 469-478. doi: 10.1534/genetics.111.132068.
- 1131 Yang, Z., Martens, C.A., Bruno, D.P., Porcella, S.F., and Moss, B. (2012). Pervasive initiation and  
1132 3'-end formation of poxvirus postreplicative RNAs. *J Biol Chem* 287(37), 31050-31060. doi:  
1133 10.1074/jbc.M112.390054.
- 1134 Yoshikawa, K., Tanaka, T., Ida, Y., Furusawa, C., Hirasawa, T., and Shimizu, H. (2011).  
1135 Comprehensive phenotypic analysis of single-gene deletion and overexpression strains of  
1136 *Saccharomyces cerevisiae*. *Yeast* 28(5), 349-361. doi: 10.1002/yea.1843.
- 1137 Zakharov, J.A., Yurchenko, L.V., and Yarovot, B.F. (1969). Cytofusion: The autonomous transfer  
1138 of cytoplasmic hereditary factors during pairing of yeast cells. *Genetika* 5, 136-141.
- 1139 Zhong, Z., Ravikumar, A., and Liu, C.C. (2018). Tunable Expression Systems for Orthogonal DNA  
1140 Replication. *ACS Synth Biol* 7(12), 2930-2934. doi: 10.1021/acssynbio.8b00400.
- 1141
- 1142
- 1143



1144 **FIGURE LEGENDS**

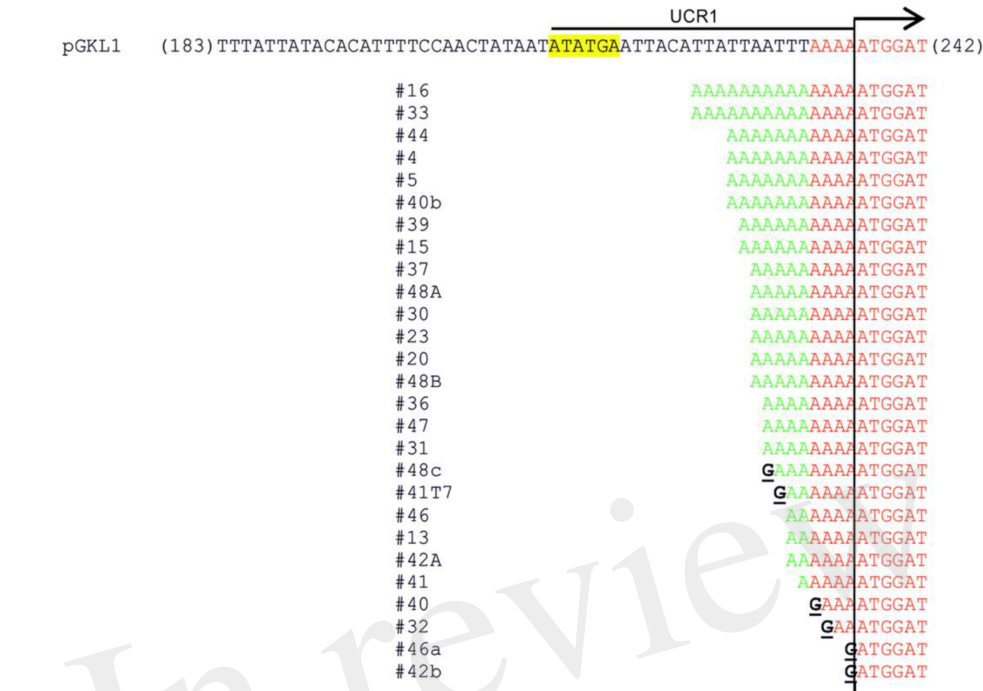


1145 **Figure 1.** Genetic organization of pGKL VLEs (linear cytoplasmic plasmids). Known gene functions  
1146 are indicated. Shades of gray indicate an extent of transcript capping as it was determined in this  
1147 work, from black ORFs (> 85% mRNAs 5' capped) to white ORFs (< 35% mRNAs 5' capped).  
1148 Transcripts from grey ORFs comprise 45% to 60% 5' capped mRNAs and represent intermediate  
1149 degree of capping. All the pGKL1 ORFs belong to the group with low occurrence of 5' capped  
1150 transcripts. TIR – terminal inverted repeat; • – terminal protein; SSB – Single-Stranded Binding  
1151 protein; TRF1 – Terminal Recognition Factor 1; ? – unknown protein function.  
1152



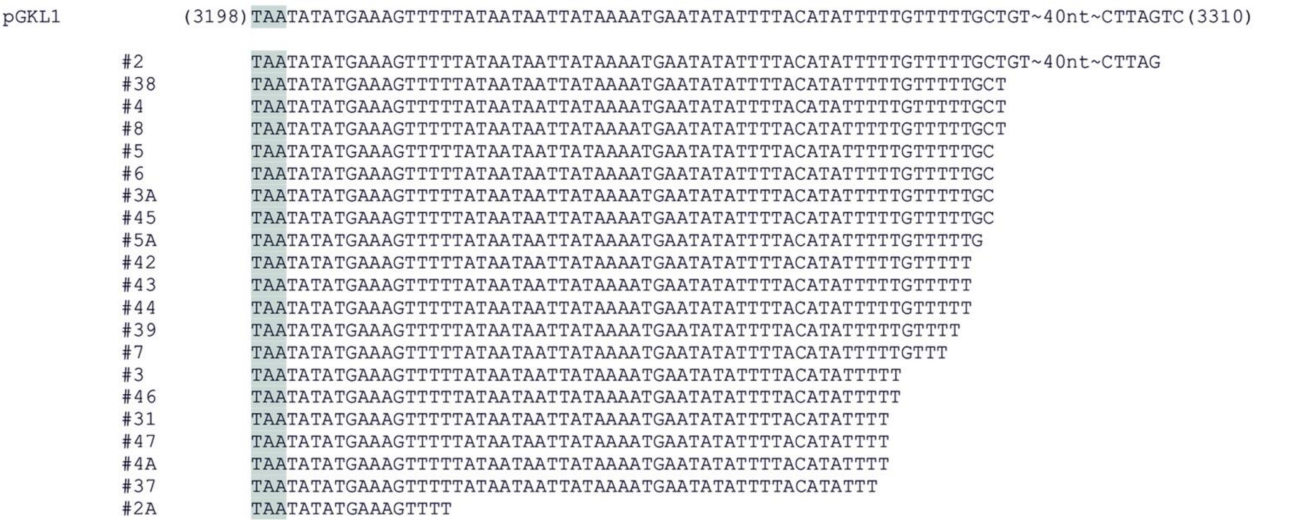
A

pGKL1-ORF1 (5' RACE)



B

pGKL1-ORF1 (3' RACE)

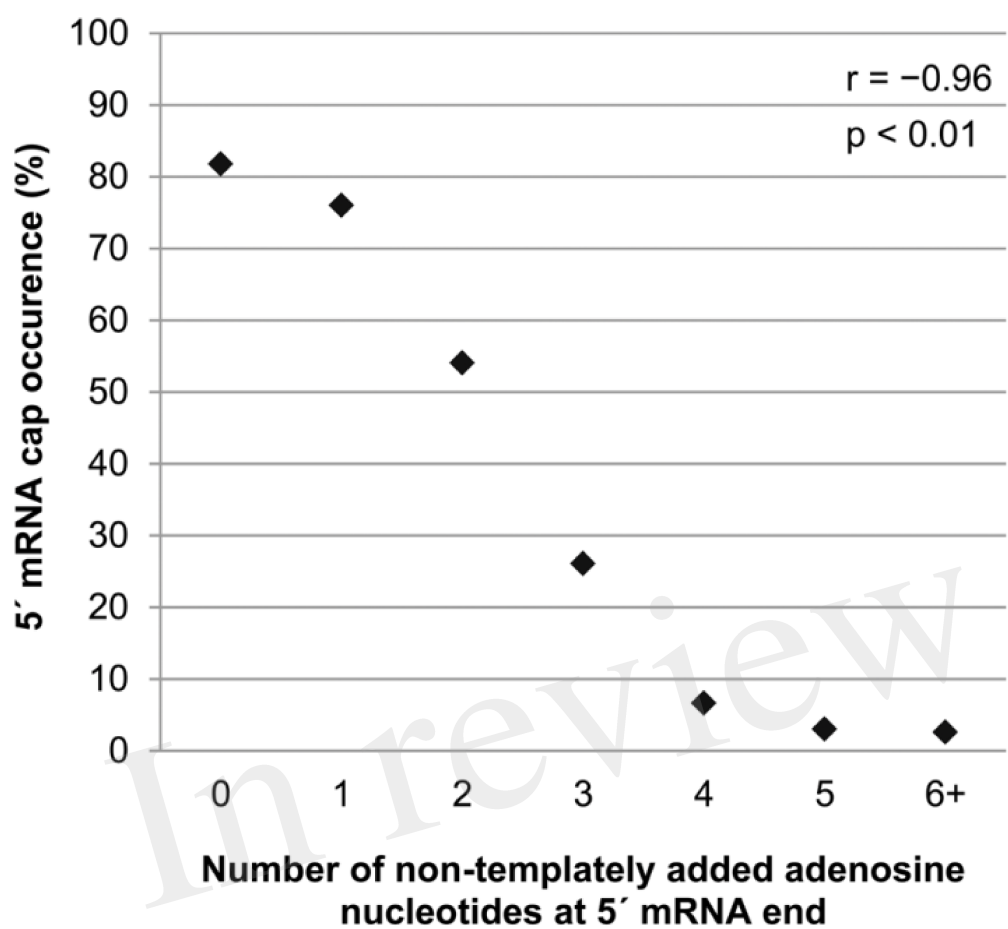


**Figure 2.** 5'/3' RACE analysis of pGKL transcripts. **(A, B)** Representative pictures of 5' RACE **(A)** and 3' RACE **(B)** analyses of individual mRNAs transcribed from *K1ORF1* encoded by the pGKL1 VLE. The top sequence corresponds to the template (VLE genomic) DNA; the cDNA sequences below represent individual clones obtained by the reverse transcription of individual mRNA molecules. In the case of 5' RACE **(A)**, sequences homologous to the template (VLE genomic) DNA

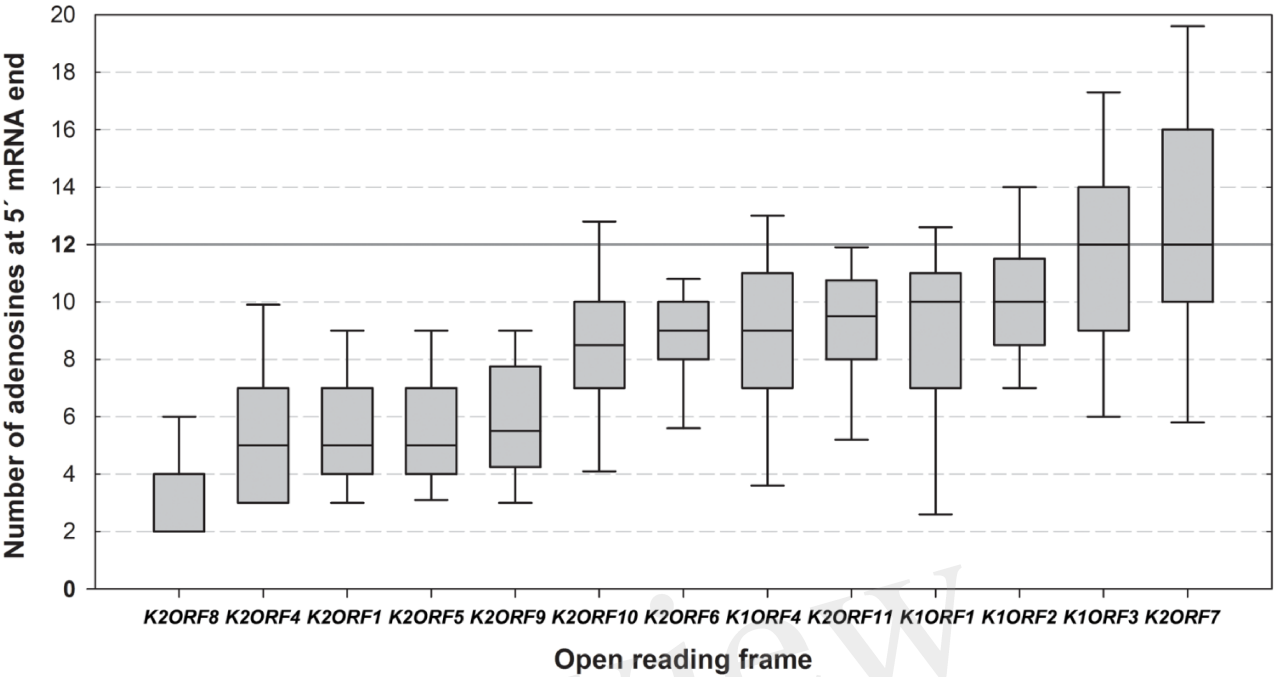
1159 are highlighted in red, and non-template 5' leaders of the mRNA transcripts are in green. Guanosine  
1160 nucleotides uncovered as cytosines during the first strand cDNA synthesis and corresponding to the  
1161 5' mRNA caps are depicted in black and underlined. *KIUCRI*, which was also used as a promoter  
1162 sequence in other experiments, is underlined. *KIUCSI* is highlighted in yellow. **(B)** Sequences  
1163 obtained by 3' RACE. Stop codons are highlighted in gray. For details concerning transcripts of other  
1164 pGKL ORFs, please refer to Table S5A and C, and the description therein.

1165

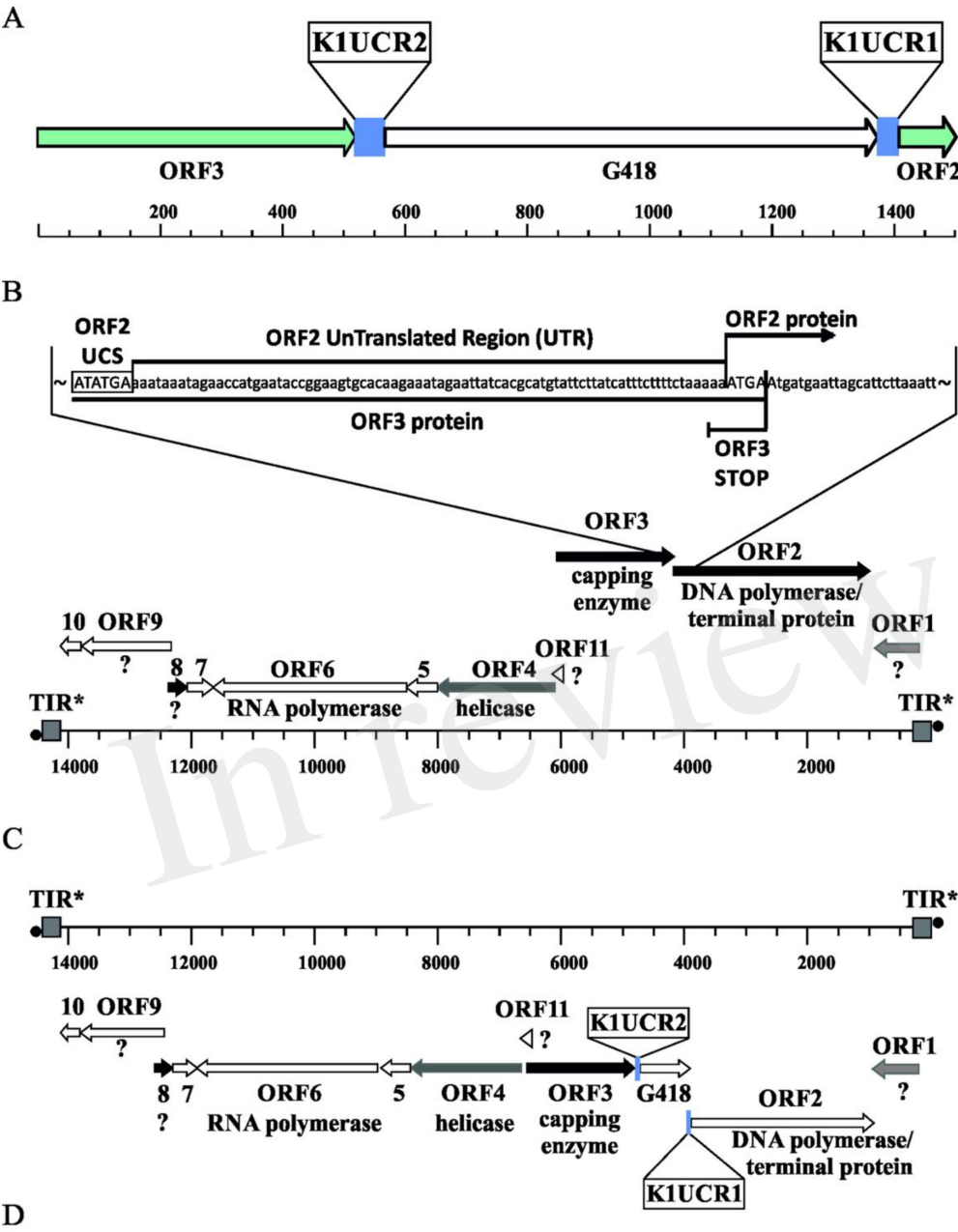
In review



**Figure 3.** Degree of non-templated 5' mRNA polyadenylation negatively correlates with 5' mRNA cap occurrence frequency in pGKL transcripts. Transcripts were divided into individual groups (Table S6) based on their number of non-templately added adenosine nucleotides in the 5' poly(A) leader. A plot represents percentage of sequences containing 5' mRNA cap for each group. Number of sequences containing 5' mRNA cap in transcripts with six and more non-templated adenosine nucleotides was very low (Table S6), therefore, these transcripts were combined into a single group. Pearson correlation coefficient value ( $r$ ) is shown. Results are statistically significant at  $p < 0.01$  level. In total, 350 sequences obtained by 5' RACE-PCR were used for this analysis.



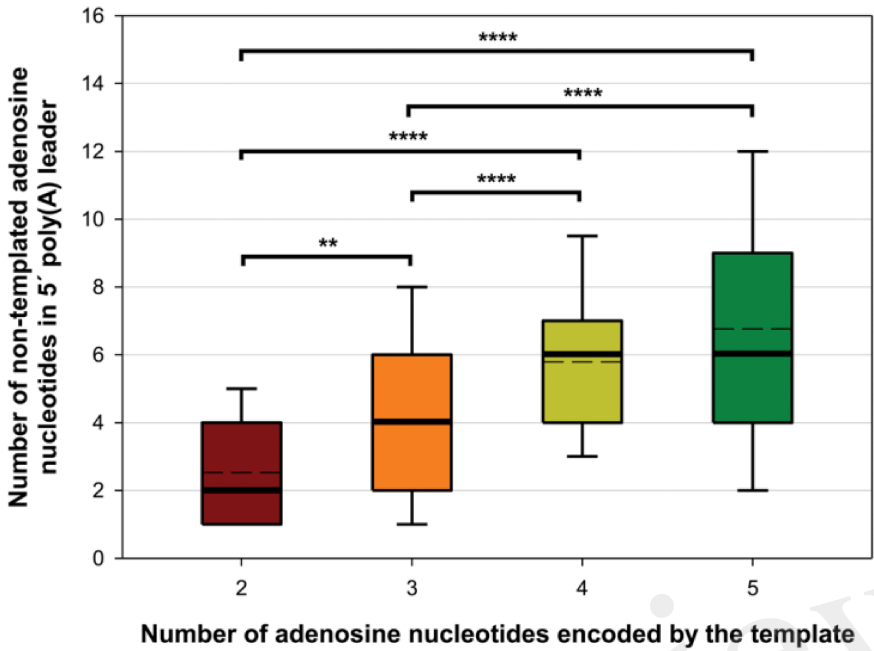
**Figure 4.** Differences in the lengths of the 5' mRNA poly(A) leaders of individual pGKL ORFs. A box whisker plot represents number of templated and non-templated consecutive adenosine nucleotides at 5' ends of pGKL mRNAs. The bottom and top of the box indicate the first and third quartile, respectively. Whiskers indicate 10th and 90th percentile. Median is indicated as a solid black line. Outliers are not indicated. The open reading frames are ranked from left to right according the prevalent length of 5' poly(A) leaders of their mRNAs. The value of 12 adenosine nucleotides represents an optimal length of the Pab1 binding site and is indicated as a solid gray line. The pGKL1 mRNAs belong to those with the longest 5' poly(A) leaders. The data did not follow normal distribution according to the Shapiro-Wilk test. Results were statistically analyzed using non-parametric Kruskal-Wallis test which supported rejection of the null hypothesis,  $p\text{-value} = 2.172940\text{e-}48$ . To discern gene pairs, transcripts of which significantly differed in lengths of their 5' poly(A) leaders, we performed a Dunn *post-hoc* tests followed by adjustment of p-values according to the Benjamini-Hochberg FDR method. Adjusted p-values corresponding to all possible gene pairs are depicted in Table S10. In total, 458 sequences obtained by 5' RACE-PCR were used for this analysis (Table S5A).



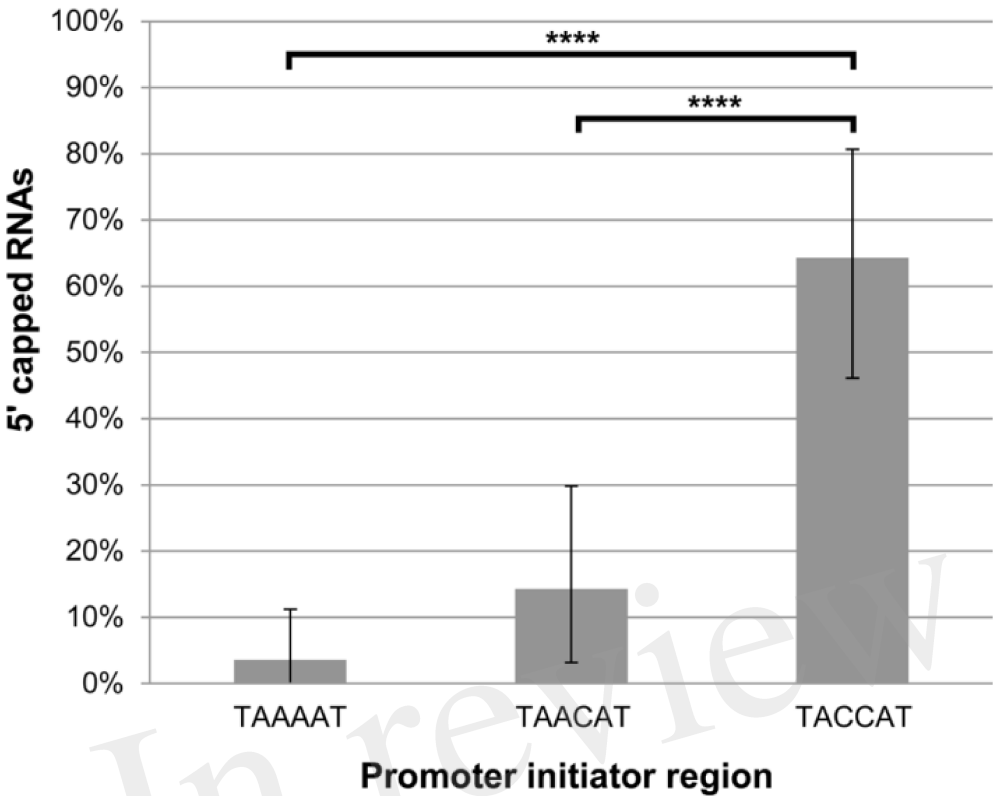
**Figure 5.** Precise manipulation of pGKL VLEs *in vivo* revealed an essential role of pGKL promoters in mRNA capping and non-template-based 5' polyadenylation. **(A)** General description of the experiment. PCR cassette used for the manipulation of pGKL2 by homologous recombination *in vivo* consists of regions homologous to the native pGKL2 VLE (in light green), an antibiotic resistance gene (*G418*) under the control of the *ORF2* promoter from pGKL1 (*K1UCR2*) and the *ORF1* promoter from pGKL1 (*K1UCR1*), which will artificially control the expression of *K2ORF2*. Promoters (UCRs) are in blue. **(B)** A closer view of the native pGKL2 region subjected to homologous recombination shows a tightly packed VLE genome. The 3' end of the *K2ORF3* coding region overlaps the *K2ORF2* promoter, 5' UTR and first 4 nucleotides of the *K2ORF2* coding region. pGKL2 VLEs displayed in panels B and C are in the reverse orientation of those in Figure 1. Shades of gray indicate the degree of transcript capping, as in Figure 1. **(C)** The resulting VLE, pRKL2-1, contains two genes, aminoglycoside 3'-phosphotransferase (coding for G418 resistance) and *K2ORF2*, that are artificially controlled by the pGKL1 promoters *K1UCR2* and *K1UCR1*, respectively. Shades of gray indicate the degree of transcript capping, as in Figure 1. **(D)** Electrophoretic analysis of pGKL VLEs in *K. lactis* clones. M - Lambda DNA/Eco130I (StyI) Marker (Fermentas); Lanes 1 and 4 - native pGKL VLEs from *K. lactis* IFO1267 (pGKL1 [8874 bp] is labelled with an asterisk, and pGKL2 [13447 bp] is labelled with an arrow). Lane 2 - linear VLEs purified from *K. lactis* IFO1267 carrying both the recombinant (higher MW) and wild type pGKL2 VLEs; Lane 3 - linear VLEs purified from *K. lactis* IFO1267 containing the recombinant pRKL2-1 VLE (14353 bp). The shorter wild type pGKL2 was lost after cultivation for  $\approx 60$  generations in selective medium containing G418.

1215

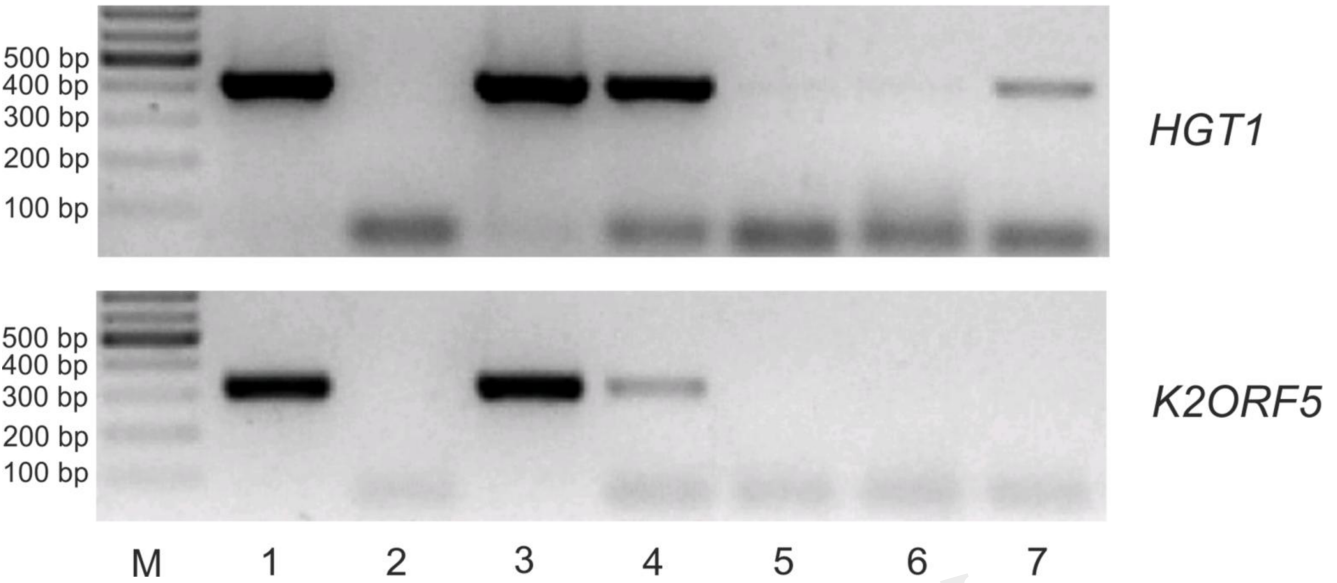




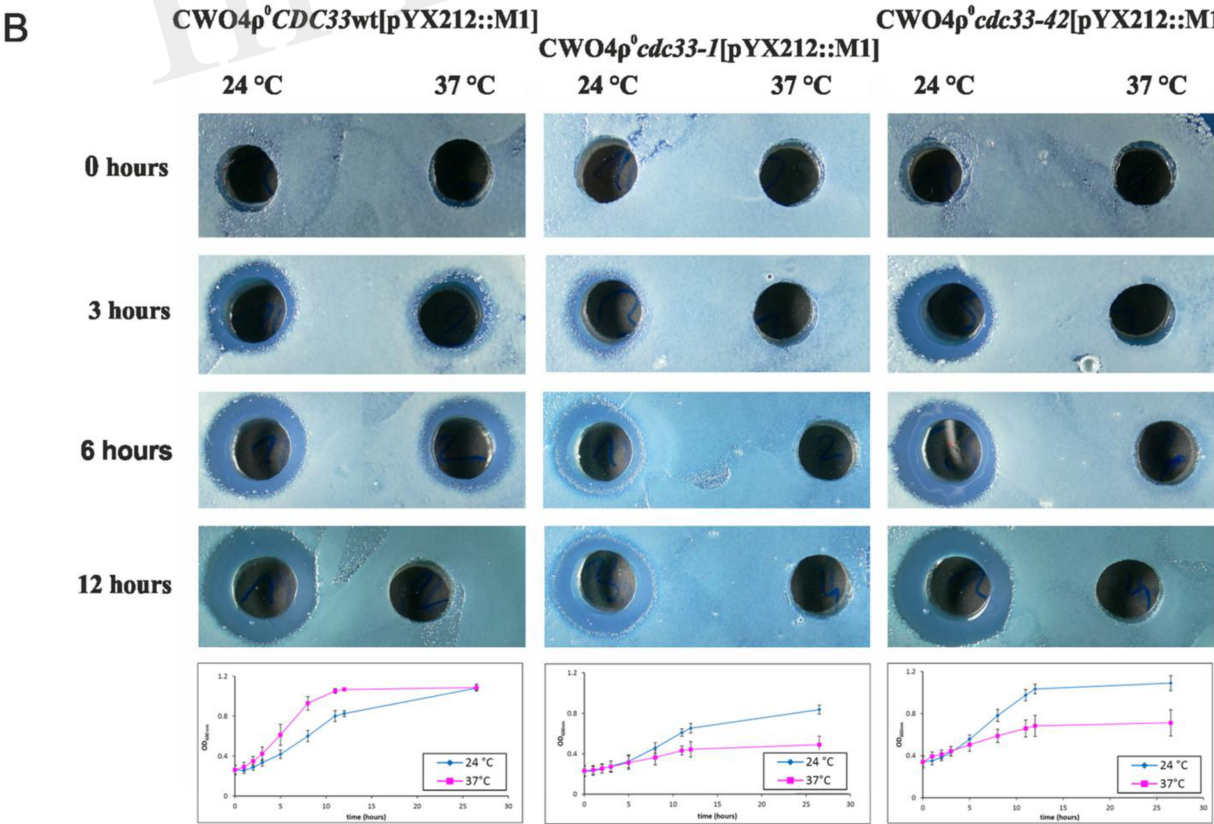
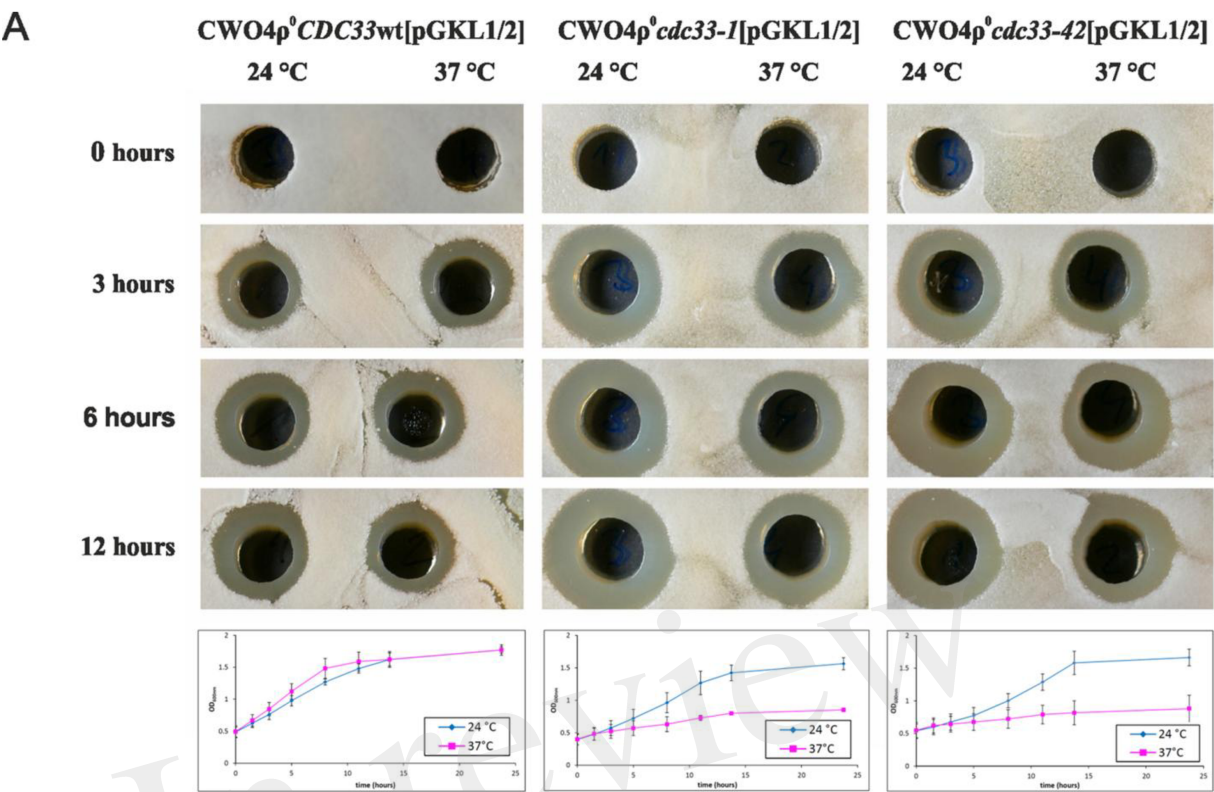
**Figure 6.** Number of 5' non-templated adenosine nucleotides in 5' leaders of pGKL transcripts increases with number of adjacent consecutive adenosines encoded by the template. A box whisker plot represents number of the non-templated adenosine nucleotides in the 5' poly(A) leaders of pGKL mRNAs classified according the length of the adjacent consecutive adenosines encoded by the DNA template at TSS. The bottom and top of the box indicate the first and third quartiles, respectively. Whiskers indicate 10th and 90th percentile. Median and mean is indicated as a bold and dashed line, respectively. Outliers are not indicated. Two adenosines are encoded in TSS of *K2ORF8* promoter (red). Three adenosines are encoded in TSS of *K1ORF4*, *K2ORF1*, *K2ORF5*, *K2ORF9*, and *K2ORF11* promoters (orange). Four adenosines are encoded in TSS of *K1ORF2*, *K2ORF6*, and *K2ORF10* promoters (yellow). Five adenosines are encoded in TSS of *K1ORF1*, *K1ORF3*, *K2ORF4*, and *K2ORF7* promoters (green). Results suggest pGKL RNA polymerase slippage at TSS as a mechanism of 5' mRNA poly(A) leader synthesis. The data did not follow normal distribution according to the Shapiro-Wilk test. Results were statistically analyzed using non-parametric Kruskal-Wallis test followed by *post-hoc* Dunn test with p-value adjustment according to the Benjamini-Hochberg FDR method. Adjusted p-values corresponding to all possible pairs are depicted in Table S11. \*\*: significance level  $p < 0.01$ . \*\*\*\*: significance level  $p < 0.0001$ . In total, 373 sequences obtained by 5' RACE-PCR containing non-templated adenosine nucleotides were used for this analysis (Table S8).



**Figure 7.** Length of the 5' poly(A) leaders of pGKL mRNAs inversely correlates with proportion of their 5' capping. The presence of 5' RNA caps was monitored using 5' RACE of the *G418<sup>R</sup>* gene from the IFO1267\_pRKL1-1 strain (TAAAAT initiator region), IFO1267\_pRKL1-2 strain (TAACAT initiator region), and IFO1267\_pRKL1-3 strain (TACCAT initiator region), followed by cloning and sequencing of individual cDNA clones (26 clones for each strain). Bars represent the frequency (in %) of 5' RNA capping with the error bars depicting the 95% confidence intervals calculated using the adjusted Wald method. Results were statistically evaluated using two-tailed Fisher's exact test with 95% confidence interval. \*\*\*\*: significance level  $p < 0.0001$ .



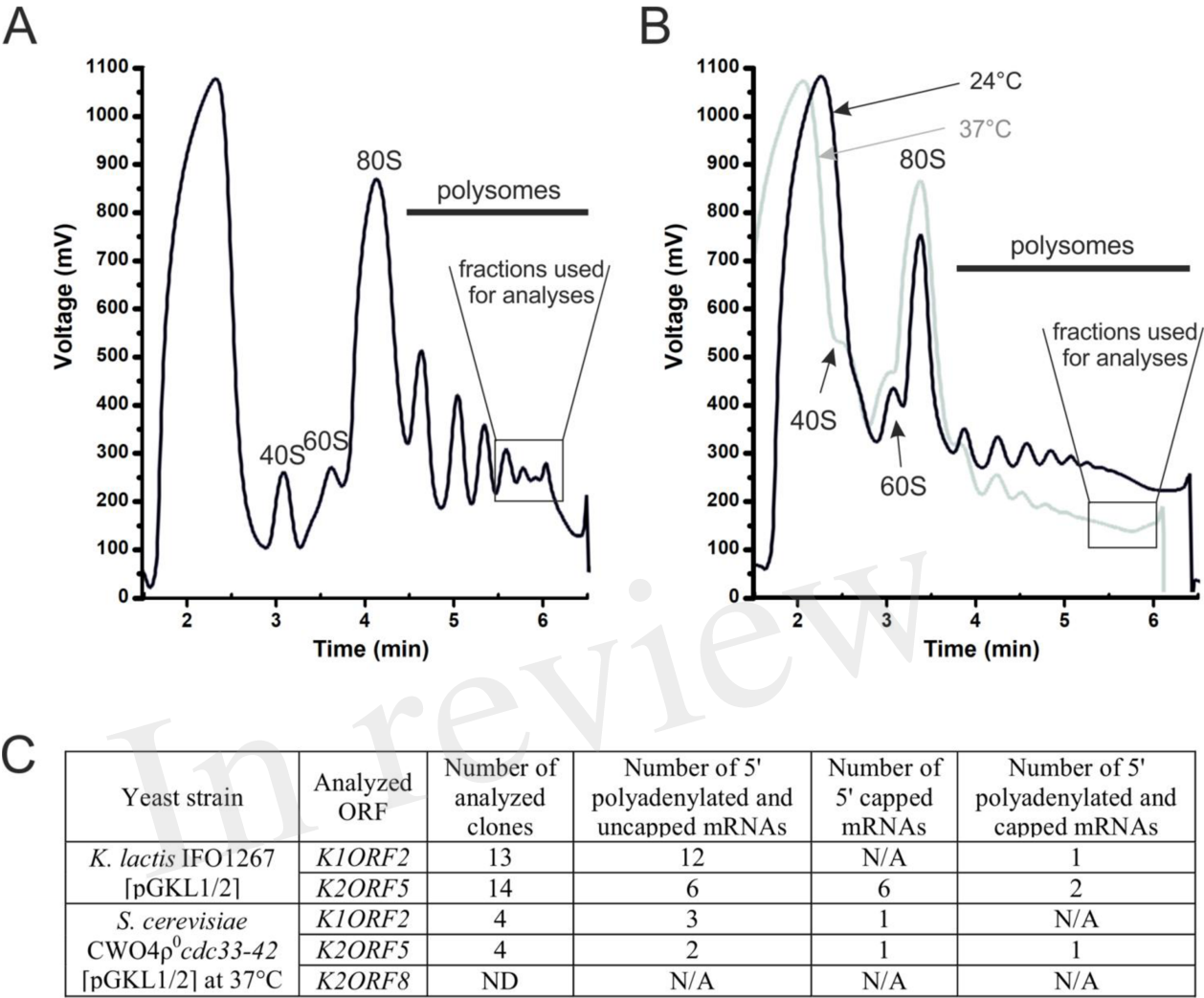
**Figure 8.** *K2ORF5* mRNA is outcompeted by cellular RNA in binding to the yeast cap-binding protein eIF4E *in vitro*. Electrophoretic analysis of semi-quantitative RT-PCR detecting control *HGT1* mRNA (upper panel) and *K2ORF5* mRNA (lower panel) in samples as follows: Lane 1 - glutathione-Sepharose with bound GST-eIF4E fusion protein in the presence of excess of *K. lactis* IFO1267 total RNA (input); Lane 2 - same as in line 1, but reaction was performed without reverse transcriptase (negative control); Lane 3 - supernatant from the first washing step (unbound mRNA); Lanes 4, 5, and 6 - supernatant after the second, third and sixth washing steps, respectively (unbound mRNA); Lane 7 - mRNA remained bound on GST-S.c-eIF4E Sepharose after the sixth washing step. M - GeneRuler 100 bp DNA Ladder Plus (Thermo Scientific). PCR was performed using cDNA, Taq DNA polymerase and gene-specific primers listed in Table S1. All washing steps were performed with 70 volumes of the buffer I. Initial abundance of *HGT1* and *K2ORF5* mRNA in the *K. lactis* total RNA were comparable as determined by qRT-PCR (Figure S3).



1260

**Figure 9.** Killer toxin encoded by pGKL VLEs is synthesized by the eIF4E-independent pathway. (A) Production of pGKL killer toxin in *S. cerevisiae* CWO4p<sup>0</sup> strains bearing cytoplasmic pGKL VLEs and different *CDC33* alleles. The pGKL toxin-coding mRNAs do not contain 3' poly(A) tails and are mostly 5' uncapped and 5' polyadenylated (Table S4 and Table S5A, Figure 6). (B) Production of K1 killer toxin in control *S. cerevisiae* CWO4p<sup>0</sup> strains containing pYX212::M1 plasmid and different *CDC33* alleles. The control mRNA coding for K1 toxin is transcribed in the nucleus by RNA polymerase II. In both cases, the results are shown for *S. cerevisiae* CWO4p<sup>0</sup> strains bearing either the wild-type *CDC33* (eIF4E) gene (*CDC33*wt) or temperature-sensitive mutations in the *CDC33* (eIF4E) gene (*cdc33-1* and *cdc33-42* alleles) grown under both permissive (24°C) and non-permissive (37°C) conditions. Aliquots were taken at 0, 3, 6 and 12 hours, and the culture medium was filter-sterilized, diluted 10x, and assayed for the presence of the killer toxin activity by an agar well diffusion test using a lawn of *S. cerevisiae* S6/1 sensitive strain cells. Growth curves of all strains at permissive (blue) and non-permissive (purple) temperatures are shown below each of the corresponding sets of killer toxin experiments. Each growth curve represents at least three independent experiments. Error bars correspond to standard deviation.





**Figure 10.** 5' polyadenylated and uncapped pGKL mRNAs are loaded to yeast polysomes independently of eIF4E translation initiation factor. (A) Analysis of polysomes from *K. lactis* IFO1267 grown exponentially in YPD medium at 28 °C. Fraction of high MW polysomes used for the mRNA purification and 5' RACE analysis of pGKL mRNAs is depicted. (B) Analysis of polysomes from the *S. cerevisiae* CWO4 *cdc33-1*  $\rho^0$  [pGKL1/2] strain grown in YPD medium in parallel at permissive (24°C, black curve) and restrictive (37°C, grey curve) temperature. Overall cellular translation is substantially decreased in absence of functional eIF4E (grey curve). Fractions of remaining high MW polysomes from culture grown at non-permissive 37°C temperature, which were used for the mRNA purification and 5' RACE analysis of pGKL mRNAs, are depicted. (C) Tabular summary of the 5' RACE results from mRNA purified from high MW polysomes in experiments A and B. Messenger RNAs of each ORF subjected to 5' RACE analysis are

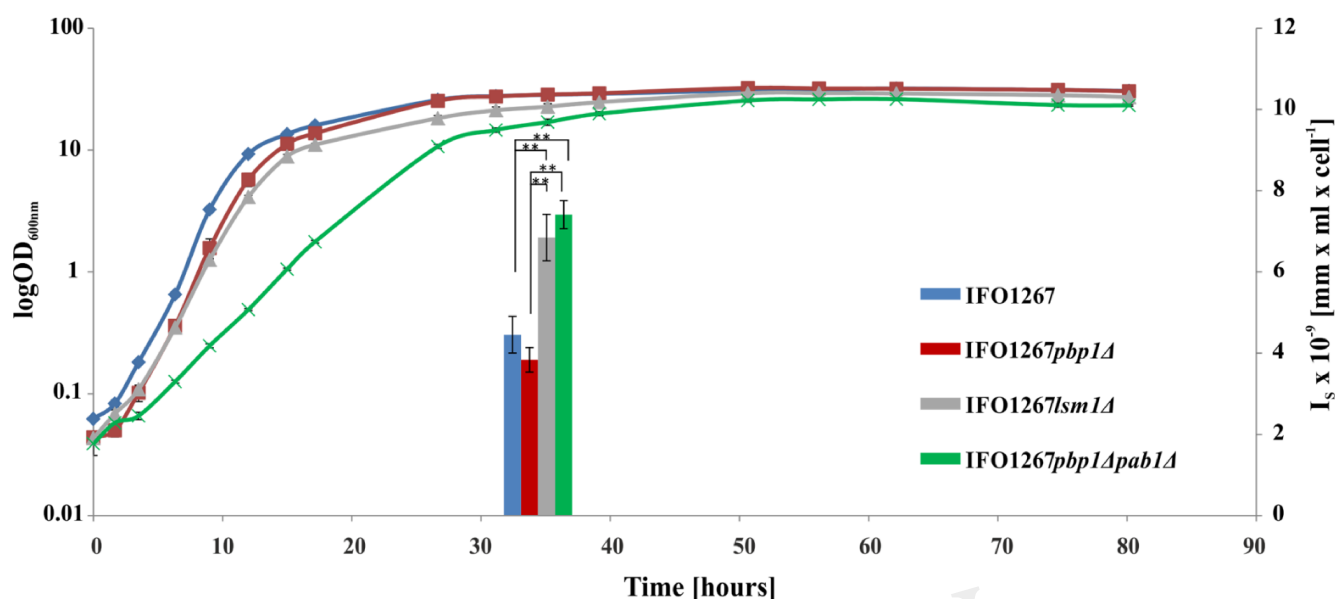


## Transcriptome of the yeast virus-like elements

1289 characterized by the following data: 1/ ORF name; 2/ total number of analysed clones; 3/ number of  
1290 5' polyadenylated and uncapped mRNAs; 4/ number of 5' capped and not 5' polyadenylated mRNAs;  
1291 5/ number of 5' polyadenylated and capped mRNAs. ND – Not Detected; N/A - Not Applicable; 40S  
1292 - eukaryotic small ribosomal subunit; 60S - eukaryotic large ribosomal subunit; 80S - eukaryotic  
1293 ribosome (monosome).

1294

In review



**Figure 11.** Translation of the *K. lactis* killer toxin is enhanced in *LSM1* and *PAB1* deletion strains. Growth curves and production of the pGKL1 killer toxin from *K. lactis* IFO1267 (blue), IFO1267 *pbp1Δ* (red), IFO1267 *lsm1Δ* (grey) and IFO1267 *pbp1Δ pab1Δ* (green) strains. All strains were grown in YPD medium at 28°C for approximately 80 hours, and their OD<sub>600 nm</sub> and toxin production were monitored. Production of the pGKL1 killer toxin into culture medium was assayed by a well diffusion test on YPD agar plates with a lawn of the *S. cerevisiae* S6/1 sensitive strain (for details of the method refer to Materials and Methods and Figure S4). Toxin levels in a culture medium in early stationary phase (~35 hours), when all the cultures reached comparable OD<sub>600</sub>, are depicted. Amount of active toxin produced by each strain is presented as a calculated normalized inhibition zone at killer toxin saturation ( $I_s$ ). Growth curves and  $I_s$  determination corresponds to at least three independent replicates. Results were statistically evaluated using one-way ANOVA followed by *post-hoc* Tukey's HSD test and further confirmed with Scheffé multiple comparison. (\*\*,  $p < 0.01$ ) Normal distribution of the data was confirmed by the Shapiro-Wilk test. Error bars at each point of the growth curves and in a bar graph represent standard deviation.